

Large deviations for rough path lifts of Watanabe's pullbacks of delta functions *

Yuzuru Inahama

Abstract

We study Donsker-Watanabe's delta functions associated with strongly hypoelliptic diffusion processes indexed by a small parameter. They are finite Borel measures on the Wiener space and admit a rough path lift. Our main result is a large deviation principle of Schilder type for the lifted measures on the geometric rough path space as the scale parameter tends to zero. As a corollary, we obtain a large deviation principle conjectured by Takanobu and Watanabe, which is a generalization of a large deviation principle of Freidlin-Wentzell type for pinned diffusion processes.

1 Introduction

In 1993 Takanobu and Watanabe [25] presented a large deviation principle (LDP) of Freidlin-Wentzell type for solutions of stochastic differential equations (SDEs) under the strong Hörmander condition anywhere. Unlike in the usual LDP of this type, the probability measures in [25] are not the push-forwards of the (scaled) Wiener measure, but the push-forwards of the measures of finite energy which is defined by the composition of the solutions of SDEs and the delta functions (i.e., Watanabe's pullbacks of the delta functions, also known as Donsker's delta function). One interpretation of this LDP is a generalization of the LDP of Freidlin-Wentzell type for pinned diffusion measures. This LDP (Theorem 2.1, [25]) looks very nice. To the author's knowledge, however, no proof has been given yet.

In this paper we reformulate this LDP on the geometric rough path space by lifting these measures in the rough path sense and prove it rigorously by using quasi-sure analysis (which is a kind of potential theory in Malliavin calculus). Then, Theorem 2.1, [25] is a simple corollary of our main result. After suitably specializing it, we also obtain the LDP for pinned diffusion measures under the strong Hörmander condition anywhere. Our main tools are rough path theory, Watanabe's distributional Malliavin calculus, and quasi-sure analysis.

***Mathematics Subject Classification:** 60F10, 60H07, 60H99, 60J60. **Keywords:** large deviation principle, rough path theory, Malliavin calculus, quasi-sure analysis, pinned diffusion process.

The elliptic case was already done in the author's previous work [15]. This work is a generalization of it to the strongly hypoelliptic case. Note that many basic results on quasi-sure analysis for Brownian rough path were already obtained in [15]. Compared to [15], the lower estimate becomes more difficult, while the upper estimate remains somewhat similar.

Let us briefly recall the history of LDP of Schilder type on rough path space. The first result was for the law of (scaled) Brownian rough path by Ledoux, Qian, and Zhang [19]. By the continuity of Lyons-Itô map in the rough path setting, the usual LDP of Freidlin-Wentzell type is immediate. Although a few nice proofs of this LDP are known, this new proof is quite straight forward and looks powerful. Since then, LDPs of Schilder type became one of the central topics in the probabilistic aspects of rough path theory and many papers have been written on it. (For example, an LDP for a wide class of Gaussian rough paths is proved in Friz and Victoir [8]. This class includes fractional Brownian rough path with Hurst parameter $H \in (1/4, 1/2]$. The original motivation of [15] was to extend the idea in [19] to the case of pinned diffusion measures.) Another advantage of this method is that one can also prove Laplace approximation (i.e., the precise asymptotics of LDP of Freidlin-Wentzell type) along the same streamline with or without Malliavin calculus. (For example, see [14, 16] for results for fractional Brownian rough path). In short, LDP theory on rough path space turned out to be quite successful. Therefore, we believe that the geometric rough path space is the right place for the LDP conjectured by Takanobu and Watanabe in [25].

The organization of this paper is as follows. In Section 2 we introduce the setting, make basic assumptions, and state our main result (Theorem 2.1) and its corollaries (Corollaries 2.2 and 2.4). Section 3 is devoted to calculations of the skeleton ODE. We prove that the deterministic Malliavin covariance matrix is non-degenerate at sufficiently many Cameron-Martin paths. This is a key lemma in the proof of the lower estimate of our main theorem. In Section 4 we present some preliminaries on quasi-sure analysis on rough path space, all of which were already shown or used in [15].

We prove the large deviation lower estimate in Section 5. Compared to the elliptic case in [15], this part becomes more difficult for two reasons. (These are closely related, however.) One is non-degeneracy of the deterministic Malliavin covariance matrix. It fails at some Cameron-Martin paths in the hypoelliptic case. (The aim of Section 3 is to deal with this difficulty). The other is that uniform non-degeneracy of Malliavin covariance matrix of the diffusion processes does not hold in general. We will use a modified version of the asymptotic theory, which turns out to fit very well with the localization procedure on the geometric rough path space with Besov type topology.

In Section 6 we prove the large deviation upper estimate. This part is not very different from the corresponding part of [15]. (However, it is not so easy for those who are not familiar with Watanabe's distributional Malliavin calculus). The key point is the integration by parts formula for Watanabe distributions, combined with Kusuoka-Stroock's quantitative proof of non-degeneracy of Malliavin covariance matrix. In Section

7, using Lyons' continuity theorem and the contraction principles for LDPs, we prove the LDP conjectured in [25] as a simple corollary of our main theorem.

2 Setting and Main results

In this section we introduce our setting and state our main results. Although the setting may seem complicated at first sight, we believe that the reader will gradually find it quite natural. The SDEs we consider in this paper and our assumptions on the coefficient vector fields are standard. Our explanation in this section may not be so detailed, but we will give precise definitions and detailed explanations in later sections.

Let $\mathcal{W} = C_0([0, 1], \mathbb{R}^d)$ be the set of the continuous functions from $[0, 1]$ to \mathbb{R}^d which start at 0. This is equipped with the usual sup-norm. We denote by \mathcal{H} and μ the Cameron-Martin subspace of \mathcal{W} and the Wiener measure on \mathcal{W} , respectively. The triple $(\mathcal{W}, \mathcal{H}, \mu)$ is called the classical Wiener space. The canonical realization of Brownian motion is denoted by $(w_t)_{0 \leq t \leq 1}$.

Let $V_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a vector field with sufficient regularity ($0 \leq i \leq d$). Precisely, we assume the following regularity condition. We say that V_i ($0 \leq i \leq d$) satisfies Assumption **(A1)** if

(A1): V_i is of C^∞ with bounded derivatives of all order ≥ 1 .

Note that V_i itself may have linear growth in the above condition. When V_i is also bounded, V_i is said to be of C_b^∞ . (C_b^k is similarly defined for $k = 1, 2, \dots$).

Let $\varepsilon \in (0, 1]$ be a small parameter. Under **(A1)**, we consider the following SDE of Stratonovich type:

$$dX_t^\varepsilon = \varepsilon \sum_{i=1}^d V_i(X_t^\varepsilon) \circ dw_t^i + \varepsilon^2 V_0(X_t^\varepsilon) dt \quad \text{with} \quad X_0^\varepsilon = x \in \mathbb{R}^n. \quad (2.1)$$

When necessary, we will write $X_t^\varepsilon = X^\varepsilon(t, x, w)$ or $X^\varepsilon(t, x)$ and sometimes write $\lambda_t^\varepsilon = \varepsilon^2 t$. Recall that **(A1)** is a standard assumption in Malliavin calculus, under which X_t^ε is \mathbf{D}_∞ -functional for all $t \geq 0$ and $\varepsilon \in (0, 1]$.

Next we impose a non-degeneracy assumption on the vector fields. We set

$$\Sigma_1 = \{V_i \mid 1 \leq i \leq d\} \quad \text{and} \quad \Sigma_k = \{[V_i, W] \mid 1 \leq i \leq d, W \in \Sigma_{k-1}\}$$

for $k \geq 2$ recursively. For $x \in \mathbb{R}^n$, we write $\Sigma_k(x) = \{W(x) \mid W \in \Sigma_k\}$, which is a finite subset of $\mathbb{R}^n \cong T_x \mathbb{R}^n$ (i.e., the tangent space at x). We assume the following strong Hörmander condition everywhere.

(A2): For any $x \in \mathbb{R}^n$, $\cup_{k=1}^\infty \Sigma_k(x)$ spans $\mathbb{R}^n \cong T_x \mathbb{R}^n$ in the sense of linear algebra.

Note that the drift vector field V_0 is not involved in **(A2)**.

In this paragraph, we will assume $t > 0$ and $\varepsilon \in (0, 1]$. It is well-known that under **(A1)**–**(A2)**, X_t^ε is non-degenerate in the sense of Malliavin. Hence, the composition $T(X_t^\varepsilon) = T \circ X_t^\varepsilon$ is well-defined as a Watanabe distribution (i.e., a generalized Wiener functional) on \mathcal{W} for any tempered Schwartz distribution T on \mathbb{R}^n . In particular, the heat kernel $p_t^\varepsilon(x, x')$ (i.e., the density of the law of $X^\varepsilon(t, x)$ with respect the Lebesgue measure dx') exists and is equal to $\mathbb{E}[\delta_{x'}(X^\varepsilon(t, x))]$, where $\delta_{x'}(X^\varepsilon(t, x))$ is Watanabe's pullback of the delta function and \mathbb{E} stands for the generalized expectation. It is known that $p_t^\varepsilon(x, x') > 0$ for all $x, x' \in \mathbb{R}^n$. (To check this positivity under our assumptions **(A1)**–**(A2)**, combine Theorem 3.41, Aida, Kusuoka, and Stroock [2] and Theorem 5.3, Kunita [17] for example.)

Now we introduce the skeleton ODE which corresponds to SDE (2.1). For a Cameron-Martin path $h \in \mathcal{H}$, we consider the following controlled ODE.

$$d\phi_t = \sum_{i=1}^d V_i(\phi_t) dh_t^i \quad \text{with} \quad \phi_0 = x \in \mathbb{R}^n. \quad (2.2)$$

Note that this ODE has a unique global solution for any given h under **(A1)**. The solution will often be denoted by $\phi_t(h)$, $\phi(t, x, h)$, etc. Note the absence of the drift term in (2.2).

Let \mathcal{V} be an l -dimensional linear subspace of \mathbb{R}^n ($1 \leq l \leq n$) and $\Pi_{\mathcal{V}} : \mathbb{R}^n \rightarrow \mathcal{V}$ be the orthogonal projection. (For our purpose, we may and sometimes will assume without loss of generality that $\mathcal{V} = \mathbb{R}^l \times \{\mathbf{0}_{n-l}\}$, where $\mathbf{0}_{n-l}$ is the zero vector of \mathbb{R}^{n-l} .) Set $Y_t^\varepsilon = \Pi_{\mathcal{V}}(X_t^\varepsilon)$, which will often be denoted by $Y^\varepsilon(t, x, w)$, and $\psi(t, x, h) = \Pi_{\mathcal{V}}\phi(t, x, h)$, where ϕ is the solution of ODE (2.2). For $x, x' \in \mathbb{R}^n$ and $a \in \mathcal{V}$, define $\mathcal{K}^{x, x'} = \{h \in \mathcal{H} \mid \phi(1, x, h) = x'\}$ and

$$\mathcal{M}^{x, a} = \{h \in \mathcal{H} \mid \psi(1, x, h) = a\} = \bigcup \{\mathcal{K}^{x, x'} \mid x' \in \Pi_{\mathcal{V}}^{-1}(a)\}.$$

By the controllability of ODE (2.2) under **(A2)** (and **(A1)**), $\mathcal{K}^{x, x'} \neq \emptyset$ for any x, x' . (See Theorem 5.3, Kunita [17]) Hence, $\mathcal{M}^{x, a} \neq \emptyset$ for any x, a .

Let $\lambda(X_t^\varepsilon)$ and $\lambda(Y_t^\varepsilon)$ be the smallest eigenvalue of the Malliavin covariance matrix of X_t^ε and Y_t^ε , respectively. It is known that $\lambda(X_t^\varepsilon)^{-1}$ has moments of all order (See Nualart [22] for example. This is in fact stronger than non-degeneracy of X_t^ε in the sense of Malliavin). Since $\Pi_{\mathcal{V}}^*$ is an isometry, $\lambda(X_t^\varepsilon) \leq \lambda(Y_t^\varepsilon)$. Hence, $\lambda(Y_t^\varepsilon)^{-1}$ also has moments of all order and consequently Y_t^ε is non-degenerate in the sense of Malliavin. (In a similar way, non-degeneracy of the deterministic Malliavin covariance of $\phi(t, x, h)$ at $h \in \mathcal{H}$ implies that of $\psi(t, x, h)$.)

Therefore, for any $a \in \mathcal{V}$, $\delta_a(Y_t^\varepsilon)$ is a positive Watanabe distribution and equal to $(\delta_a \circ \Pi_{\mathcal{V}})(X_t^\varepsilon)$. By the positivity of $p_t^\varepsilon(x, x')$, we can easily see that $\mathbb{E}[\delta_a(Y_t^\varepsilon)] > 0$ for all $\varepsilon \in (0, 1]$, $t > 0$, $x \in \mathbb{R}^n$, and $a \in \mathcal{V}$. By Sugita's theorem [24], the positive Watanabe distribution $\delta_a(Y_1^\varepsilon)$ at time $t = 1$ is in fact a finite Borel measure on \mathcal{W} , which will be denoted by $\theta_{x, a}^\varepsilon$.

From here we discuss rough path theory. In this paper, we consider the geometric rough path space $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ with Besov-type topology. We will always assume that the Besov parameter $(\alpha, 4m)$ satisfies the following assumption so that basic results in [15] are available;

$$\frac{1}{3} < \alpha < \frac{1}{2}, \quad m = 1, 2, 3, \dots, \quad \alpha - \frac{1}{4m} > \frac{1}{3}, \quad \text{and} \quad 4m(\frac{1}{2} - \alpha) > 1. \quad (2.3)$$

Under (2.3), $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ is continuously embedded in the geometric rough path space $G\Omega_{\alpha'}^H(\mathbb{R}^d)$ with Hölder topology with $\alpha' = \alpha - 1/(4m)$. Intuitively, α is something like the Hölder exponent and $4m$ is a very large even integer.

Let $\mathcal{L} : \mathcal{W} \rightarrow G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ be the rough path lift map via the dyadic polygonal approximations, which is defined outside a slim subset of \mathcal{W} and ∞ -quasi continuous. (When this is regarded as a geometric rough path space-valued random variable, we will often write $\mathbf{W} = \mathcal{L}(w)$. On the other hand, a generic element of the geometric rough path space is denoted by \mathbf{w} , which is not random.) Since \mathcal{L} is quasi-surely defined, we can lift the measure $\theta_{x,a}^\varepsilon$ to a measure on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. We write $\mu_{x,a}^\varepsilon = (\varepsilon \cdot \mathcal{L})_*[\theta_{x,a}^\varepsilon]$, where the "dot" stands for the dilation on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. We denote by $\hat{\theta}_{x,a}^\varepsilon$ and $\hat{\mu}_{x,a}^\varepsilon$ the normalized measure of $\theta_{x,a}^\varepsilon$ and $\mu_{x,a}^\varepsilon$, respectively. (Since the total mass of $\theta_{x,a}^\varepsilon$ or of $\mu_{x,a}^\varepsilon$ equals $\mathbb{E}[\delta_a(Y_t^\varepsilon)] > 0$, this normalization is well-defined.)

Set a rate function $I_1 : G\Omega_{\alpha,4m}^B(\mathbf{R}^n) \rightarrow [0, \infty]$ as follows;

$$I_1(\mathbf{w}) = \begin{cases} \|h\|_{\mathcal{H}}^2/2 & \text{(if } \mathbf{w} = \mathcal{L}(h) \text{ for some } h \in \mathcal{M}^{x,a}), \\ \infty & \text{(otherwise).} \end{cases}$$

This rate function I_1 is actually good. We also set $\hat{I}_1(\mathbf{w}) = I_1(\mathbf{w}) - \min\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{M}^{x,a}\}$. Note that the minimum above exists.

The following theorem is our main result in this paper. It states that the family of finite measures $\{\mu_{x,a}^\varepsilon\}_{0 < \varepsilon \leq 1}$ satisfies an LDP of Schilder type on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ as $\varepsilon \searrow 0$.

Theorem 2.1 *Assume (A1) and (A2) and the condition (2.3). Then, we have the following (i)–(ii):*

(i) *The family $\{\mu_{x,a}^\varepsilon\}_{\varepsilon > 0}$ of finite measures satisfies an LDP on $G\Omega_{\alpha,4m}^B(\mathbf{R}^d)$ as $\varepsilon \searrow 0$ with a good rate function I_1 , that is, for any Borel set $A \subset G\Omega_{\alpha,4m}^B(\mathbf{R}^d)$, the following inequalities hold;*

$$-\inf_{\mathbf{w} \in A^\circ} I_1(\mathbf{w}) \leq \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) \leq \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) \leq -\inf_{\mathbf{w} \in \bar{A}} I_1(\mathbf{w}).$$

(ii) *The family $\{\hat{\mu}_{a,a'}^\varepsilon\}_{\varepsilon > 0}$ of probability measures satisfies an LDP on $G\Omega_{\alpha,4m}^B(\mathbf{R}^d)$ as $\varepsilon \searrow 0$ with a good rate function \hat{I}_1 .*

Since the whole set is both open and closed, Theorem 2.1, (i) implies that

$$\lim_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(G\Omega_{\alpha,4m}^B(\mathbf{R}^d)) = \lim_{\varepsilon \searrow 0} \varepsilon^2 \log \mathbb{E}[\delta_a(Y_t^\varepsilon)] = -\min\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{M}^{x,a}\}.$$

Therefore, Theorem 2.1, (ii) is immediate from (i).

Theorem 2.1 above also holds with respect to α' -Hölder geometric rough path topology for any $\alpha' \in (1/3, 1/2)$, because we can find α, m with (2.3) such that $(\alpha, 4m)$ -Besov topology is stronger than α' -Hölder topology.

From the contraction principle for LDPs, it is obvious that, for any continuous map F from the geometric rough path space to a Hausdorff topological space, the image measure $F_*[\hat{\mu}_{x,a}^\varepsilon]$ satisfies an LDP, too. As an example of such continuous maps, we may take a Lyons-Itô map associated with coefficient vector fields which are different from V_i 's.

Let $A_i : \mathbb{R}^N \rightarrow \mathbb{R}^N$ be a vector field which satisfies **(A1)** ($0 \leq i \leq d$). Note that they may be different from V_i . For $\varepsilon \in (0, 1]$, we also consider the following SDE of Stratonovich type;

$$dZ_t^\varepsilon = \varepsilon \sum_{i=1}^d A_i(Z_t^\varepsilon) \circ dw_t^i + \varepsilon^2 A_0(Z_t^\varepsilon) dt \quad \text{with} \quad Z_0^\varepsilon = z \in \mathbb{R}^N. \quad (2.4)$$

For $h \in \mathcal{H}$, we consider the following controlled ODE;

$$d\zeta_t = \sum_{i=1}^d A_i(\zeta_t) dh_t^i \quad \text{with} \quad \zeta_0 = z \in \mathbb{R}^N. \quad (2.5)$$

We may write $Z_t^\varepsilon = Z^\varepsilon(t, z, w)$ or $\zeta_t = \zeta_t(h) = \zeta(t, z, h)$, etc. We denote by $\tilde{Z}^\varepsilon = \tilde{Z}^\varepsilon(\cdot, z, w)$ an ∞ -quasi continuous modification of

$$\mathcal{W} \ni w \mapsto Z^\varepsilon(\cdot, z, w) \in C^{\alpha-H}([0, 1], \mathbb{R}^N) \quad (1/3 < \alpha < 1/2). \quad (2.6)$$

Here, the set on the right hand side stands for the space of α -Hölder continuous paths in \mathbb{R}^N . Since \tilde{Z}^ε is defined uniquely up to a slim subset of \mathcal{W} , the pushforward measures of $\theta_{x,a}^\varepsilon$ and $\hat{\theta}_{x,a}^\varepsilon$ by the map \tilde{Z}^ε are well-defined.

As a corollary of our main theorem, we can prove an LDP as $\varepsilon \searrow 0$ for these measures. Before stating it, let us first define good rate functions $I_2, \hat{I}_2 : C^{\alpha-H}([0, 1], \mathbb{R}^N) \rightarrow [0, \infty]$. Set

$$I_2(b) = \begin{cases} \inf\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{M}^{x,a} \text{ such that } b = \zeta(\cdot, z, h)\}, \\ \infty, & (\text{if no } h \in \mathcal{M}^{x,a} \text{ satisfies that } b = \zeta(\cdot, z, h)). \end{cases}$$

and $\hat{I}_2(b) = I_2(b) - \min\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{M}^{x,a}\}$.

Corollary 2.2 *Let $1/3 < \alpha < 1/2$. Assume **(A1)** for both V_i and A_i ($0 \leq i \leq d$) and assume **(A2)** for V_i ($0 \leq i \leq d$). Then, we have the following (i)–(ii):*

(i) *The family $\{\tilde{Z}^\varepsilon(\cdot, z)_*[\theta_{x,a}^\varepsilon]\}_{\varepsilon>0}$ satisfies an LDP on $C^{\alpha-H}([0, 1], \mathbb{R}^N)$ as $\varepsilon \searrow 0$ with a good rate function I_2 , that is, for any Borel set $A \subset C^{\alpha-H}([0, 1], \mathbb{R}^N)$, the following inequalities hold;*

$$\begin{aligned} - \inf_{b \in A^\circ} I_2(b) &\leq \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \theta_{x,a}^\varepsilon(\{w \in \mathcal{W} \mid \tilde{Z}^\varepsilon(\cdot, z, w) \in A\}) \\ &\leq \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \theta_{x,a}^\varepsilon(\{w \in \mathcal{W} \mid \tilde{Z}^\varepsilon(\cdot, z, w) \in A\}) \leq - \inf_{b \in A} I_2(b). \end{aligned}$$

(ii) The family $\{\tilde{Z}^\varepsilon(\cdot, z)_*[\hat{\theta}_{x,a}^\varepsilon]\}_{\varepsilon>0}$ of probability measures satisfies an LDP on $C^{\alpha-H}([0, 1], \mathbb{R}^N)$ as $\varepsilon \searrow 0$ with a good rate function \hat{I}_2 .

Remark 2.3 In the formulation of Corollary 2.2 above, Hölder path space $C^{\alpha-H}([0, 1], \mathbb{R}^N)$ is used, while Besov-type path space is used in Theorem 2.1, p. 200, [25]. However, by adjusting Hölder/Besov parameters, we see that these two formulations are equivalent. Therefore, Corollary 2.2 above is equivalent to the LDP conjectured in Theorem 2.1, [25].

Corollary 2.2 above immediately implies an LDP of Freidlin-Wentzell type for pinned diffusion measures as follows.

Take $n = l = N$, $x = z$ and $V_i = A_i$ for all i . We write $a = x' \in \mathbb{R}^n$. Then, $X_t^\varepsilon = Y_t^\varepsilon = Z_t^\varepsilon$, $\phi_t = \psi_t = \zeta_t$, and $\mathcal{M}^{x,a} = \mathcal{K}^{x,x'}$. In this case, $\tilde{Z}^\varepsilon(\cdot, z)_*[\hat{\theta}_{x,a}^\varepsilon]$ is nothing but the pinned diffusion measure $Q_{x,x'}^\varepsilon$ associated to the generator $\varepsilon^2\{V_0 + (1/2)\sum_{i=1}^d V_i^2\}$ (or equivalently, to the heat kernel p_t^ε) with the starting point x and the ending point x' .

Then, we have the following result. The proof is almost obvious.

Corollary 2.4 Let $1/3 < \alpha < 1/2$ and assume **(A1)** and **(A2)**. The family $\{Q_{x,x'}^\varepsilon\}_{\varepsilon>0}$ satisfies an LDP on $C^{\alpha-H}([0, 1], \mathbb{R}^N)$ as $\varepsilon \searrow 0$ with a good rate function I'_2 . Here, I'_2 is given by

$$I'_2(b) = \begin{cases} \inf\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{K}^{x,x'} \text{ such that } b = \phi(\cdot, z, h)\} - \min\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{K}^{x,x'}\}, \\ \infty, & (\text{if no } h \in \mathcal{K}^{x,x'} \text{ satisfies that } b = \phi(\cdot, z, h)). \end{cases}$$

We remark that Bailleul proved an LDP parallel to Corollary 2.4 on compact manifolds in [3] (and in its extended version [4] with Mesnager and Norris). Their method is basically analytic (with a little bit of rough path theory) and different from ours. Their result can be viewed as a hypoelliptic version of Hsu's result in [10] for pinned Brownian motions on compact Riemannian manifolds.

Remark 2.5 One cannot replace the "strong Hörmander" condition in Theorem 2.1 by the "Hörmander" condition. We have the following counterexample. Consider the following two-dimensional SDE driven by one-dimensional Brownian motion.

$$dX_t^{\varepsilon,1} = \varepsilon dw_t, \quad dX_t^{\varepsilon,2} = \varepsilon^2 X_t^{\varepsilon,1} dt.$$

The coefficient vector fields satisfy the Hörmander condition everywhere, but nowhere the strong Hörmander condition. If the solution starts at the origin, the law of $(X_1^{\varepsilon,1}, X_1^{\varepsilon,2})$ is the centered Gaussian measure with the covariance

$$\begin{pmatrix} \varepsilon^2 & \varepsilon^4/2 \\ \varepsilon^4/2 & \varepsilon^6/3 \end{pmatrix}.$$

Then, it is easy to see that $p_1^\varepsilon((0,0), (0,x^2)) = \sqrt{3}(\pi\varepsilon^4)^{-1} \exp(-(x^2)^2/(6\varepsilon^6))$. If $x^2 \neq 0$, then $\lim_{\varepsilon \searrow 0} \varepsilon^2 \log p_1^\varepsilon((0,0), (0,x^2)) = -\infty$. On the other hand, we have $\mathcal{K}^{(0,0),(0,x^2)} \neq \emptyset$. Therefore, the heat kernel does not behave in the way described in Theorem 2.1. (Recall that the heat kernel is the weight of the whole set in our setting).

Remark 2.6 *Loosely speaking, our main results above generalize the ones for the elliptic case in the author's previous paper [15]. However, the results in this paper do not cover all of the results in [15] for the following reasons:*

- (i) *In this paper the strong Hörmander condition is assumed at any point, while in [15], the ellipticity condition is assumed only at the starting point and at some point vector fields may even be degenerate (i.e., do not even satisfy the Hörmander condition).*
- (ii) *In [15] the drift vector field is of the form $V_0(\varepsilon, x)$ and is quite general. However, it is of the form $\varepsilon^2 V_0(x)$ in this paper. Although it may be possible to generalize our results for a drift term of the form $V_0(\varepsilon, x)$ with $V_0(0, x) \equiv 0$, it is probably impossible if $V_0(0, x)$ do not vanish identically. (This guess is based on an observation of small noise asymptotics of the heat kernel in Section 3, Ben Arous and Léandre [5]).*

3 Skeleton ODE

In this section we study the solution $\phi_t(h) = \phi(t, x, h)$ of the skeleton ODE (2.2). Note that it always has a global solution under **(A1)**. The aim of this section is to prove that a Fréchet differentiable map $h \mapsto \phi(1, x, h)$ is non-degenerate at sufficiently many h 's under strong Hörmander condition on the vector fields. (See Proposition 3.1). It will play a crucial role in the lower estimate for the LDP in our main theorem (Theorem 2.1). We emphasize again that the absence of the drift term in (2.2) has a significant meaning and many parts of this section would fail if (2.2) had a drift term.

3.1 Basic properties of skeleton ODE

First we set some notations. For $T > 0$, \mathcal{H}_T denotes \mathbb{R}^d -valued Cameron-Martin space on the time interval $[0, T]$, that is,

$$\mathcal{H}_T = \{h : [0, T] \rightarrow \mathbb{R}^d \mid h = \int_0^\cdot \dot{h}_s ds \text{ for some } \dot{h} \in L^2([0, T], \mathbb{R}^d)\}.$$

The Hilbert norm is naturally defined by $\|h\|_{\mathcal{H}_T} = \|\dot{h}\|_{L^2[0, T]}$ as usual. When $T = 1$, we simply write $\mathcal{H} = \mathcal{H}_1$.

For $h \in \mathcal{H}_T$, the reversed path $\bar{h} \in \mathcal{H}_T$ is defined by $\bar{h}_t = h_{T-t} - h_T$. Concatenation of $h \in \mathcal{H}_T$ and $k \in \mathcal{H}_S$ is denoted by $h * k \in \mathcal{H}_{T+S}$, which is defined by $(h * k)(t) = h(t)$ for $0 \leq t \leq T$ and $(h * k)(t) = k_{t-T} + h_T$ for $T \leq t \leq T + S$.

Thanks to **(A2)**, ODE (2.2) is strongly completely controllable (Theorem 5.3, Kunita [17]). Hence, for any x, x' and $T > 0$, there exists $h \in \mathcal{H}_T$ such that $\phi(T, x, h) = x'$ and, in particular, $\mathcal{K}^{x, x'} := \{h \in \mathcal{H} \mid \phi(1, x, h) = x'\} \neq \emptyset$.

Now we introduce Jacobian ODE of (2.2) and its inverse.

$$dJ_t = \sum_{i=1}^d \nabla V_i(\phi_t) J_t dh_t^i \quad \text{with} \quad J_0 = \text{Id}_n, \quad (3.1)$$

$$dK_t = - \sum_{i=1}^d K_t \nabla V_i(\phi_t) dh_t^i \quad \text{with} \quad K_0 = \text{Id}_n. \quad (3.2)$$

Here, $J, K, \nabla V_i$ are all $n \times n$ matrices. Note that $K_t = J_t^{-1}$. When dependency on h and x needs to be specified, we write $J_t(h)$ or $J(t, x, h)$, etc.

The map $h \in \mathcal{H} \mapsto \phi_t(h) = \phi(t, x, h) \in \mathbb{R}^n$ is of Fréchet- C^1 for each $t \in [0, 1]$ and $x \in \mathbb{R}^n$. The Fréchet derivative $D\phi_t(h) \in L(\mathcal{H}, \mathbb{R}^n)$ is explicitly given by

$$D\phi_t(h)\langle k \rangle = J_t(h) \sum_{i=1}^n \int_0^t J_s(h)^{-1} V_i(\phi_s(h)) \dot{k}_s^i ds \quad (3.3)$$

The deterministic Malliavin covariance of ϕ_1 at h (and at time $t = 1$) is defined by

$$\sigma_{\phi_1}(h) = \left(\langle D\phi_1^i(h), D\phi_1^j(h) \rangle_{\mathcal{H}^*} \right)_{1 \leq i, j \leq n} = D\phi_1(h) \circ [D\phi_1(h)]^*, \quad (3.4)$$

where the superscript $*$ stands for the adjoint operation. From (3.3) and (3.4) we can easily see that $\sigma_{\phi_1}(h) = J_1(h)C(h)J_1(h)^*$ with

$$C(h) = \int_0^1 J_s(h)^{-1} \mathbf{V}(\phi_s(h)) \mathbf{V}(\phi_s(h))^* J_s(h)^{-1,*} ds \quad (3.5)$$

Here, we set $\mathbf{V}(x) = [V_1(x), \dots, V_d(x)]$, which is an $n \times d$ matrix. Note that the surjectivity of the linear map $D\phi_1(h) : \mathcal{H} \rightarrow \mathbb{R}^n$ is equivalent to non-degeneracy of the deterministic Malliavin covariance $\sigma_{\phi_1}(h)$, which in turn is equivalent to non-degeneracy of $C(h)$ since $J_1(h)$ is always invertible.

The following is the main result in this section. Unlike in the elliptic case, there exists h such that $\sigma_{\phi_1}(h)$ is degenerate. (For example, think of the constant path $0 \in \mathcal{H}$.) However, there are sufficiently many h 's for which $\sigma_{\phi_1}(h)$ is non-degenerate. The precise statement is given as follows.

Proposition 3.1 *Assume (A1) and (A2). Let $x, x' \in \mathbb{R}^n$ and $h \in \mathcal{K}^{x, x'}$ be arbitrary. Then, we have the following;*

- (i) *For any $\varepsilon > 0$, there exists $h^\varepsilon \in \mathcal{K}^{x, x'}$ such that $\|h - h^\varepsilon\|_{\mathcal{H}} < \varepsilon$ and $\sigma_{\phi_1}(h^\varepsilon)$ is non-degenerate.*
- (ii) *Moreover, h^ε in (i) above can be chosen so that $\langle h^\varepsilon, \cdot \rangle_{\mathcal{H}}$ naturally extends to a continuous linear functional on the Wiener space \mathcal{W} .*

The proof of Proposition 3.1, (i) will be given in the subsequent subsections. Once we have Proposition 3.1, (i), we can prove (ii) by using the following lemma (with $\mathcal{K} = \mathcal{H}$, $\mathcal{L} = \mathcal{W}^*$ and $\mathcal{W}^* \subset \mathcal{H}^* \cong \mathcal{H}$).

Lemma 3.2 *Let \mathcal{K} be a real Hilbert space and $\xi \in \mathcal{K}$. Assume that (i) F is an \mathbb{R}^n -valued Fréchet- C^1 map defined on a neighborhood of ξ with a bounded derivative DF and (ii) $DF(\xi) : \mathcal{K} \rightarrow \mathbb{R}^n$ is a surjective linear map. Let \mathcal{L} be a real Banach space which is continuously and densely embedded in \mathcal{K} . Then, there exists $\xi_j \in \mathcal{L}$ ($j = 1, 2, \dots$) such that $\lim_{j \rightarrow \infty} \|\xi_j - \xi\|_{\mathcal{K}} = 0$ and $F(\xi_j) = F(\xi)$ for all j . (Necessarily, $DF(\xi_j)$ is also surjective for large enough j .)*

Proof. This lemma was proved in [15]. ■

Before closing this subsection, we prove two simple lemmas for later use. For $h \in \mathcal{H}$ and $x \in \mathbb{R}^n$ and a vector field $W : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we set $Q_t^W = J_t^{-1}W(\phi_t)$. Note that both J^{-1} and ϕ depend on h and x . (We will sometimes write $Q_t^W(h)$ or $Q^W(t, x, h)$, etc.)

Lemma 3.3 (i) *Let $W : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a smooth vector field. Then, we have*

$$dQ_t^W(h) = \sum_{i=1}^d Q_t^{[V_i, W]}(h) \dot{h}_t^i dt.$$

(ii) *For any $v \in \mathbb{R}^n$, we have*

$$v^*C(h)v = \sum_{i=1}^d \int_0^1 |\langle v, Q_s^{V_i}(h) \rangle|^2 ds.$$

In particular, if for any v with $\|v\| = 1$, there exist $t \in [0, 1]$ and i ($1 \leq i \leq d$) such that $\langle v, Q_t^{V_i}(h) \rangle \neq 0$, then $C(h)$ and $\sigma_{\phi_1}(h)$ are non-degenerate.

Proof. The first assertion can easily be seen from (2.2)–(3.2). The second one is shown by simple calculation of block matrices and is a routine. So, the proof is omitted. ■

The next lemma is quite simple. (So we omit a proof.) However, note that the absence of a drift term in (2.2)–(3.2) is crucially important here. If they had a drift term, this lemma would fail.

Lemma 3.4 (i) *If $(\phi_t, J_t, K_t)_{0 \leq t \leq T}$ is the solution of ODEs (2.2), (3.1), and (3.2) driven by $h \in \mathcal{H}_T$ with the initial condition $(x, \text{Id}_n, \text{Id}_n)$, then $(\phi_{T-t}, J_{T-t}, K_{T-t})_{0 \leq t \leq T}$ is the solution of (2.2), (3.1), and (3.2) driven by the reversed path $\bar{h} \in \mathcal{H}_T$, with the initial condition (ϕ_T, J_T, K_T)*

(ii) *For $h \in \mathcal{H}_T$, let $(\phi_t, J_t, K_t)_{0 \leq t \leq 2T}$ be the solution of (2.2), (3.1), and (3.2) driven by $h * \bar{h} \in \mathcal{H}_{2T}$, then $(\phi_{2T}, J_{2T}, K_{2T}) = (\phi_0, J_0, K_0)$.*

(iii) *Let $T > 0$ and $\beta \in (0, T)$. If $(\phi_t, J_t, K_t)_{0 \leq t \leq T}$ is the solution of (2.2), (3.1), and (3.2) driven by $h \in \mathcal{H}_T$ with the initial condition $(x, \text{Id}_n, \text{Id}_n)$, then*

$$(\phi_{T(t-\beta)/(T-\beta)}, J_{T(t-\beta)/(T-\beta)}, K_{T(t-\beta)/(T-\beta)})_{\beta \leq t \leq T}$$

is the solution of (2.2), (3.1), and (3.2) driven by k with the initial condition $(\phi_\beta, J_\beta, K_\beta) = (x, \text{Id}_n, \text{Id}_n)$, where k is defined by $k_t = h_{T(t-\beta)/(T-\beta)}$ on the time interval $[\beta, T]$.

Lemma 3.5 *Let $h \in \mathcal{H}$. Assume that, for any $\beta \in (0, 1)$, $h^\beta \in \mathcal{H}$ which satisfies the following property is given:*

$$|\dot{h}_t^\beta| \leq 1 \text{ for a.a. } t \in [0, \beta] \quad \text{and} \quad h_t^\beta = h_{(t-\beta)/(1-\beta)} \text{ on } [\beta, 1].$$

Then, $h^\beta \rightarrow h$ in \mathcal{H} as $\beta \searrow 0$.

Proof. Without loss of generality, we may assume $d = 1$. It is sufficient to show that $\dot{h}^\beta \rightarrow \dot{h}$ in L^2 -norm.

For any $\varepsilon > 0$, there exists a continuous function $f : [0, 1] \rightarrow \mathbb{R}$ such that $\|f - \dot{h}\|_{L^2} < \varepsilon$. It is easy to see that

$$\|\dot{h} - \dot{h}^\beta\|_{L^2} \leq \|\dot{h}\|_{L^2[0, \beta]} + \|\dot{h}^\beta\|_{L^2[0, \beta]} + \|\dot{h} - \dot{h}^\beta\|_{L^2[\beta, 1]}.$$

The first and the second terms clearly vanish as $\beta \searrow 0$. The third term is dominated by

$$\|\dot{h} - f\|_{L^2[\beta, 1]} + \left\| f - \frac{1}{1-\beta} f\left(\frac{\cdot - \beta}{1-\beta}\right) \right\|_{L^2[\beta, 1]} + \frac{1}{1-\beta} \left\| f\left(\frac{\cdot - \beta}{1-\beta}\right) - \dot{h}\left(\frac{\cdot - \beta}{1-\beta}\right) \right\|_{L^2[\beta, 1]}.$$

From the way f is chosen, the sum of the first and the third term is dominated by 3ε if $\beta > 0$ is sufficiently small. Due to the uniform continuity of f , the second term vanishes as $\beta \searrow 0$. Thus, we have shown $\limsup_{\beta \searrow 0} \|\dot{h} - \dot{h}^\beta\|_{L^2} \leq 3\varepsilon$. Letting $\varepsilon \searrow 0$, we finish the proof of the lemma. \blacksquare

We will fix an arbitrary initial point x in what follows. We say $\{V_i\}$ satisfies strong Hörmander condition of degree N at x if N is the smallest integer such that $\cup_{k=1}^N \Sigma_k(x)$ linearly spans \mathbb{R}^n . In this case there exists a subset $\Lambda \subset \cup_{k=1}^N \Sigma_k$ of cardinality n such that $\{W(x') \mid W \in \Lambda\}$ linearly spans \mathbb{R}^n for any x' which is sufficiently close to x . (We will write $\Lambda = \{W_1, \dots, W_n\}$.) By compactness, it holds that $3\lambda := \inf_{v \in \mathbb{S}^{n-1}} \max_{1 \leq j \leq n} |\langle v, W_j(x) \rangle| > 0$, where $\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n \mid |v| = 1\}$ is the unit sphere.

Lemma 3.6 *Keep the same notations as above. For sufficiently small $T > 0$, the following property holds: For any $v \in \mathbb{S}^{n-1}$, there exists $W \in \Lambda$ such that*

$$\inf\{|\langle v, Q_t^W(h) \rangle| \mid 0 \leq t \leq T, \quad h \in \mathcal{H} \text{ with } |\dot{h}_s| \leq 1 \text{ for a.a. } s \in [0, T]\} \geq \lambda > 0.$$

Proof. Set $E_j = \{v \in \mathbb{S}^{n-1} \mid |\langle v, W_j(x) \rangle| \geq 3\lambda\}$ for $1 \leq j \leq n$. Then, each E_j is compact and $\mathbb{S}^{n-1} = \cup_j E_j$. Since W_j is continuous in x , there exists $r > 0$ such that

$$\inf\{|\langle v, W_j(x') \rangle| \mid x' \in B_r(x), v \in E_j\} \geq 2\lambda > 0 \quad (1 \leq j \leq n),$$

where $B_r(x) = \{x' \in \mathbb{R}^n \mid |x' - x| < r\}$ is the ball of radius $r > 0$ centered at x .

Let $v \in E_j$. If T is sufficiently small, then $\phi_t(h)$ stays inside $B_r(x)$. Therefore, $|\langle v, W_j(\phi_t(h)) \rangle| \geq 2\lambda$ when $0 \leq t \leq T$. On the other hand, there exists a constant $M > 0$ (independent of such an h) which satisfies that $|J_t(h)^{-1} - J_s(h)^{-1}| \leq M|t - s|$ for all $s, t \in [0, T]$. Hence, by taking $T \leq \lambda M^{-1}(\max_j \sup\{|W_j(x')| \mid x' \in B_r(x)\})^{-1}$, we can prove the lemma since $Q_t^{W_j}(h) = J_t(h)^{-1}W_j(\phi_t(h))$. Notice that the choice of T is independent of v . \blacksquare

Remark 3.7 *In what follows, the constants $r, T \in (0, 1)$ which appear in (the proof of) Lemma 3.6 above will be fixed. (Of course, so will $\lambda > 0$.)*

3.2 Proof of Proposition 3.1 (i): For degrees 1, 2, 3

In this subsection we prove Proposition 3.1 (i) when N , i.e., the degree of hypoellipticity at the initial point x , is 1, 2 or 3. (Strictly speaking, this subsection is not necessary. However, we believe it helps the reader understand what is going on in the proof for the general case in the next subsection.)

Before doing so, we set a few notations for general N . First, Let $L > 0$ be the smallest constant such that

$$|W(x_1) - W(x_2)| \leq L|x_1 - x_2| \quad (x_1, x_2 \in B_r(x), W \in \cup_{k=1}^N \Sigma_k).$$

Note that r has already been determined. From this we can see the following estimate (3.6): Let $T > 0$ be as in Lemma 3.6. Then, there exists $M > 0$ such that for any $s, t \in [0, T]$, $W \in \cup_{k=1}^N \Sigma_k$ and h with $|\dot{h}_t| \leq 1$ for a.a. $t \in [0, T]$, it holds that

$$|Q_t^W(h) - Q_s^W(h)| \leq M|t - s|. \quad (3.6)$$

For $\tau > 0$, $i \in \{1, \dots, d\}$, and $\kappa \in \{\pm 1\}$, we set $\xi^{\tau, i, \kappa} \in \mathcal{H}_\tau$ by $\dot{\xi}_t^{\tau, i, \kappa} = \kappa \mathbf{1}_{[0, \tau]}(t) \mathbf{e}_i$, where $\mathbf{1}$ denotes the indicator function and $\{\mathbf{e}_i\}_{i=1}^d$ denotes the canonical basis of \mathbf{R}^d .

For $\tau_1, \dots, \tau_N > 0$, $i_1, \dots, i_N \in \{1, \dots, d\}$, and $\kappa_1, \dots, \kappa_N \in \{\pm 1\}$, we will consider $\xi^{\tau_1, i_1, \kappa_1} * \dots * \xi^{\tau_N, i_N, \kappa_N} \in \mathcal{H}_{\tau_1 + \dots + \tau_N}$, whose derivative in time is given by

$$\sum_{k=1}^N \kappa_k \mathbf{1}_{[\tau_1 + \dots + \tau_{k-1}, \tau_1 + \dots + \tau_k]}(t) \mathbf{e}_{i_k} \quad (0 \leq t \leq \tau_1 + \dots + \tau_N).$$

(When $k = 1$, $\tau_1 + \dots + \tau_{k-1}$ is understood to be 0.) For $h \in \mathcal{H}_\tau$ with $\tau > 0$, we will write $\mathcal{A}h = h * \bar{h} \in \mathcal{H}_{2\tau}$.

The case $N = 1$ (i.e., the elliptic case) is almost obvious, because for any h and any $v \in \mathbb{S}^{n-1}$ there is i such that $\langle v, Q_0^{V_j}(h) \rangle = \langle v, V_i(x) \rangle \neq 0$, which implies $\sigma_{\phi_1}(h)$ is non-degenerate for any h .

Next, we consider the case $N = 2$. Let $\tau \in (0, T]$, where T is the constant in Lemma 3.6. We will prove the following:

Lemma 3.8 *Let τ be as above. For any $v \in \mathbb{S}^{n-1}$, there exist i ($1 \leq i \leq d$), $\kappa \in \{\pm 1\}$ such that $\langle v, Q_t^{V_m}(\xi^{\tau, i, \kappa}) \rangle \neq 0$ for some m ($1 \leq m \leq d$) and some $t \in [0, \tau]$.*

Proof. Take any v and let $W \in \Lambda$ be as in Lemma 3.6. Since we assume $N = 2$, W is of the form either $W = V_j$ or $W = [V_j, V_k]$. If $W = V_j$ for some j , then for any i and κ , $\langle v, Q_0^{V_j}(\xi^{\tau, i, \kappa}) \rangle = \langle v, V_j(x) \rangle \neq 0$.

Suppose that $W = [V_j, V_k]$ for some j, k . If $\langle v, V_k(x) \rangle \neq 0$, then the same argument as above can still be used. So, we may assume that $\langle v, V_k(x) \rangle = 0$. Take $i = j$ and $\kappa = +1$. By Lemma 3.3, (i), we have

$$\left| \frac{d}{dt} \langle v, Q_t^{V_k}(\xi^{\tau, j, +1}) \rangle \right| = |\langle v, Q_t^{[V_j, V_k]}(\xi^{\tau, j, +1}) \rangle| \geq \lambda > 0 \quad (0 \leq t \leq \tau).$$

Here, we also used Lemma 3.6 and the choice of W . Since the initial value of $\langle v, Q_t^{V_k}(\xi^{\tau, j, +1}) \rangle$ is assumed to be 0, $\langle v, Q_t^{V_k}(\xi^{\tau, j, +1}) \rangle \neq 0$ for any small $t > 0$. \blacksquare

Lemma 3.9 *When $N = 2$, Proposition 3.1, (i) is true.*

Proof. For $\tau \in (0, T]$, consider $\mathcal{A}^{\xi^{\tau, i, \kappa}} = \xi^{\tau, i, \kappa} * \overline{\xi^{\tau, i, \kappa}}$ for all i, κ and concatenate them all, which is called k^τ . (The order of concatenation does not matter.) Since there are $2d$ such $\xi^{\tau, i, \kappa}$'s, the total times length is $2d \times 2\tau = 4d\tau$. So, $k^\tau \in \mathcal{H}_{4d\tau}$. We consider ODEs (2.2)–(3.2) driven by k^τ . By Lemma 3.4, (ii),

$$(\phi_{2\tau l}(k^\tau), J_{2\tau l}(k^\tau), J_{2\tau l}(k^\tau)^{-1}) = (x, \text{Id}_n, \text{Id}_n) \quad \text{for all } l = 0, 1, \dots, 2d.$$

This means that, at times $2\tau, 4\tau, \dots, 4d\tau$, the solution (ϕ_t, J_t, J_t^{-1}) gets back to the initial state and starts all over again. (If these ODEs had a drift term, this argument would fail.)

Set $\beta = 4d\tau$ and define h^β for a given $h \in \mathcal{K}^{x, x'}$ as follows. On $[0, \beta]$, we set $h_t^\beta = k_t^\tau$. On $[\beta, 1]$, we set $h_t^\beta = h_{(t-\beta)/(1-\beta)}$. Then, by Lemma 3.5, $h^\beta \rightarrow h$ in \mathcal{H} as $\tau \searrow 0$. Moreover, by Lemma 3.4, (iii), $h^\beta \in \mathcal{K}^{x, x'}$. By Lemma 3.8 and the way we construct k^τ (and h^β), we have the following: For any $v \in \mathbb{S}^{n-1}$, there exist j ($1 \leq j \leq d$) and $t \in [0, \beta]$ such that $\langle v, Q_t^{V_j}(h^\beta) \rangle \neq 0$. This implies non-degeneracy of $\sigma_{\phi_1}(h^\beta)$. \blacksquare

In the end of this subsection, we consider the case $N = 3$. If one understands the proof for this case, then one will easily understand the proof for the general case in the next subsection.

Lemma 3.10 *For sufficiently small $\tau > 0$, we set $\tau_1 = \tau$ and $\tau_2 = \lambda\tau/(2M)$, where $M > 0$ is a constant given in (3.6). Then, for any $v \in \mathbb{S}^{n-1}$, there exist $i_1, i_2 \in \{1, \dots, d\}$, $\kappa_1, \kappa_2 \in \{\pm 1\}$ such that $\langle v, Q_t^{V_m}(\xi^{\tau_1, i_1, \kappa_1} * \xi^{\tau_2, i_2, \kappa_2}) \rangle \neq 0$ for some m ($1 \leq m \leq d$) and some $t \in [0, \tau_1 + \tau_2]$.*

Proof. We take τ so small that $\tau_1 + \tau_2 \leq T$, where T is given in Lemma 3.6. Take any v and let $W \in \Lambda$ be as in Lemma 3.6. We assume that W is of the form $W = [V_j, [V_k, V_l]]$ since the other cases are easier.

On the first subinterval $[0, \tau_1]$, choose $i_1 = j$. On this interval $Q_t^V(\xi^{\tau_1, i_1, \kappa_1} * \xi^{\tau_2, i_2, \kappa_2}) = Q_t^V(\xi^{\tau_1, j, \kappa_1})$. By Lemma 3.3, (i), we have

$$\frac{d}{dt}\langle v, Q_t^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1}) \rangle = \kappa_1 \langle v, Q_t^{[V_j, [V_k, V_l]]}(\xi^{\tau, j, \kappa_1}) \rangle \quad (0 \leq t \leq \tau_1). \quad (3.7)$$

On this subinterval, the right hand side of (3.7) is of constant sign, due to Lemma 3.6. If the initial value $\langle v, Q_0^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1}) \rangle = \langle v, [V_k, V_l](x) \rangle \geq 0$, then we choose κ_1 so that the right hand side of (3.7) is positive. If otherwise, then we choose κ_1 so that the right hand side of (3.7) is negative. Either way, we have $|\langle v, Q_{\tau_1}^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1}) \rangle| \geq \lambda\tau$.

On the second subinterval $[\tau_1, \tau_1 + \tau_2]$, choose $i_2 = k$ and consider $\xi^{\tau_1, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}$. By (3.6) and the definition of τ_2 ,

$$|\langle v, Q_t^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle - \langle v, Q_{\tau_1}^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle| \leq \frac{\lambda\tau}{2} \quad (\tau_1 \leq t \leq \tau_1 + \tau_2).$$

Hence, we have $|\langle v, Q_t^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle| \geq \lambda\tau/2$ on the second subinterval. By Lemma 3.3, (i), we have

$$\frac{d}{dt}\langle v, Q_t^{V_l}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle = \kappa_2 \langle v, Q_t^{[V_k, V_l]}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle \quad (\tau_1 \leq t \leq \tau_1 + \tau_2). \quad (3.8)$$

If the initial value $\langle v, Q_{\tau_1}^{V_l}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle$ of this subinterval is non-negative, then we choose κ_2 so that the right hand side of (3.8) is positive. If otherwise, then we choose κ_2 so that the right hand side of (3.8) is negative. Either way,

$$|\langle v, Q_{\tau_1 + \tau_2}^{V_l}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle| \geq \tau_2 \times \frac{\lambda\tau}{2} = \frac{(\lambda\tau)^2}{4M}.$$

This completes the proof. (In fact, in order to prove this lemma it is enough to assume that $\langle v, Q_{\tau_1}^{V_l}(\xi^{\tau, j, \kappa_1} * \xi^{\tau_2, k, \kappa_2}) \rangle = 0$ above. However, we deliberately argued in this way for later use.) ■

Lemma 3.11 *When $N = 3$, Proposition 3.1, (i) is true.*

Proof. Let τ be sufficiently small and τ_1, τ_2 be as above. Consider

$$\mathcal{A}(\xi^{\tau_1, i_1, \kappa_1} * \xi^{\tau_2, i_2, \kappa_2})$$

for all $i_1, i_2, \kappa_1, \kappa_2$. (There are $(2d)^2$ of them.) The concatenation of all of them is denoted by k^τ . (The order of concatenation does not matter.) The total time length β of k^τ is given by

$$\beta = (2d)^2 \cdot 2(\tau_1 + \tau_2) = O(\tau) \quad \text{as } \tau \rightarrow 0.$$

On $[0, \beta]$, we set $h_t^\beta = k_t^\tau$. On $[\beta, 1]$, we set $h_t^\beta = h_{(t-\beta)/(1-\beta)}$. The rest is essentially the same as the proof for the case $N = 2$. ■

3.3 Proof of Proposition 3.1 (i): The general case

Now, we are in a position to prove Proposition 3.1 (i) for the general degree $N \geq 1$.

For sufficiently small $\tau > 0$, we set

$$\tau_1 = \tau \quad \text{and} \quad \tau_l = 2 \left(\frac{\lambda\tau}{4M} \right)^{2^{l-2}} \quad \text{for } 2 \leq l \leq N-1.$$

Here, $M > 0$ is a constant given in (3.6).

Lemma 3.12 *Let $\tau > 0$ be sufficiently small and set τ_l ($1 \leq l \leq N-1$) as above. Then, for any $v \in \mathbf{S}^{n-1}$, there exist $i_l \in \{1, \dots, d\}$, $\kappa_l \in \{\pm 1\}$ ($1 \leq l \leq N-1$) such that*

$$\langle v, Q_t^{V_m}(\xi^{\tau_1, i_1, \kappa_1} * \dots * \xi^{\tau_{N-1}, i_{N-1}, \kappa_{N-1}}) \rangle \neq 0$$

for some m ($1 \leq m \leq d$) and some $t \in [0, \tau_1 + \dots + \tau_{N-1}]$.

Proof. For simplicity we write $T_l := \tau_1 + \dots + \tau_l$. The proof is similar to the ones for Lemmas 3.8 and 3.10. Take any v and let $W \in \Lambda$ be as in Lemma 3.6. We assume that W is of the form $W = [V_{j_1}, \dots, [V_{j_{N-2}}, [V_{j_{N-1}}, V_{j_N}]] \dots]$ since the other cases are easier. In this case we take $i_l = j_l$ for $1 \leq l \leq N-1$ and write $\eta = \xi^{\tau_1, j_1, \kappa_1} * \dots * \xi^{\tau_{N-1}, j_{N-1}, \kappa_{N-1}}$. We will see that for a suitable choice of κ_l 's, $\langle v, Q_t^{V_m}(\eta) \rangle \neq 0$ holds for $m = j_N$ and $t = T_{N-1}$.

Write $U_l = [V_{j_l}, \dots, [V_{j_{N-1}}, V_{j_N}]] \dots]$ for $1 \leq l \leq N-1$ and $U_N = V_{j_N}$. On the l th interval ($1 \leq l \leq N-1$), we have

$$\frac{d}{dt} \langle v, Q_t^{U_{l+1}}(\eta) \rangle = \kappa_l \langle v, Q_t^{U_l}(\eta) \rangle \quad (T_{l-1} \leq t \leq T_l). \quad (3.9)$$

We will prove by induction that, for suitable choices of κ_i 's,

$$|\langle v, Q_{T_l}^{U_{l+1}}(\eta) \rangle| \geq 4M \left(\frac{\lambda\tau}{4M} \right)^{2^{l-1}} \quad \text{for all } 2 \leq l \leq N-1. \quad (3.10)$$

Once this is obtained, the proof of the lemma is done since $U_N = V_{j_N}$. In the same way as in Lemma 3.10, we can prove that (3.10) holds for $l = 2$ for a suitable choice of κ_1 and κ_2 . Let us assume that (3.10) holds up to $l-1$ for some $\kappa_1, \dots, \kappa_{l-1}$. By the Lipschitz continuity (3.6) and the definition of τ_l ,

$$|\langle v, Q_t^{U_l}(\eta) \rangle - \langle v, Q_{T_{l-1}}^{U_l}(\eta) \rangle| \leq M\tau_l = 2M \left(\frac{\lambda\tau}{4M} \right)^{2^{l-2}} \quad (T_{l-1} \leq t \leq T_l).$$

From this estimate and (3.10) with $l-1$,

$$|\langle v, Q_t^{U_l}(\eta) \rangle| \geq 2M \left(\frac{\lambda\tau}{4M} \right)^{2^{l-2}} \quad (T_{l-1} \leq t \leq T_l).$$

Hence, the right hand side of (3.9) is of constant sign. If $\langle v, Q_{T_{l-1}}^{U_{l+1}}(\eta) \rangle$ is non-negative (or non-positive), then we choose $\kappa_l = \pm 1$ so that the right hand side of (3.9) is positive (or negative, respectively). Then, it follows that

$$|\langle v, Q_{T_l}^{U_{l+1}}(\eta) \rangle| \geq 2M \left(\frac{\lambda\tau}{4M} \right)^{2^{l-2}} \tau_l = 4M \left(\frac{\lambda\tau}{4M} \right)^{2^{l-1}},$$

which shows that (3.10) holds up to l . Thus, we have proved (3.10). \blacksquare

Lemma 3.13 *Proposition 3.1, (i) is true for any $N \geq 1$.*

Proof. Once we obtain Lemma 3.12 above, the proof of the lemma is similar to that of Lemma 3.11. Let τ be sufficiently small and $\tau_1, \dots, \tau_{N-1}$ be as above. Consider

$$\mathcal{A}(\xi^{\tau_1, i_1, \kappa_1} * \dots * \xi^{\tau_{N-1}, i_{N-1}, \kappa_{N-1}})$$

for all i_l, κ_l ($1 \leq l \leq N-1$) and concatenate them all. (The order of concatenation does not matter.) The total time length $\beta := 2T_{N-1}(2d)^{N-1}$ is clearly of $O(\tau)$ as $\tau \searrow 0$. The rest is the same as the proof for the case $N = 3$ in Lemma 3.11. \blacksquare

4 Preliminaries

4.1 Preliminaries from Malliavin calculus

We first recall Watanabe's theory of generalized Wiener functionals (i.e., Watanabe distributions) in Malliavin calculus. Most of the contents and the notations in this section are borrowed from Sections V.8–V.10, Ikeda and Watanabe [11] with trivial modifications. There is no new result in this section. Shigekawa [23] and Nualart [22] are also good textbooks of Malliavin calculus. For basic results of quasi-sure analysis, we refer to Chapter II, Malliavin [21].

Let $(\mathcal{W}, \mathcal{H}, \mu)$ be the classical Wiener space as before. (The results in this subsection also hold on any abstract Wiener space, however.) The following are of particular importance in this paper:

- (a) Basics of Sobolev spaces $\mathbf{D}_{p,r}(\mathcal{K})$ of \mathcal{K} -valued (generalized) Wiener functionals, where $p \in (1, \infty)$, $r \in \mathbb{R}$, and \mathcal{K} is a real separable Hilbert space. As usual, we will use the spaces $\mathbf{D}_\infty(\mathcal{K}) = \cap_{k=1}^\infty \cap_{1 < p < \infty} \mathbf{D}_{p,k}(\mathcal{K})$, $\tilde{\mathbf{D}}_\infty(\mathcal{K}) = \cap_{k=1}^\infty \cup_{1 < p < \infty} \mathbf{D}_{p,k}(\mathcal{K})$ of test functionals and the spaces $\mathbf{D}_{-\infty}(\mathcal{K}) = \cup_{k=1}^\infty \cup_{1 < p < \infty} \mathbf{D}_{p,-k}(\mathcal{K})$, $\tilde{\mathbf{D}}_{-\infty}(\mathcal{K}) = \cup_{k=1}^\infty \cap_{1 < p < \infty} \mathbf{D}_{p,-k}(\mathcal{K})$ of Watanabe distributions as in [11]. When $\mathcal{K} = \mathbb{R}$, we simply write $\mathbf{D}_{p,r}$, etc.
- (b) Meyer's equivalence of Sobolev norms. (Theorem 8.4, [11]. A stronger version can

be found in Theorem 4.6, [23])

(c) Pullback $T \circ F = T(F) \in \tilde{\mathbf{D}}_{-\infty}$ of tempered Schwartz distribution $T \in \mathcal{S}'(\mathbb{R}^n)$ on \mathbb{R}^n by a non-degenerate Wiener functional $F \in \mathbf{D}_{\infty}(\mathbb{R}^n)$. (see Sections 5.9, [11].)

(d) A generalized version of integration by parts formula in the sense of Malliavin calculus for Watanabe distribution, which is given as follows (See p. 377, [11]):

For a non-degenerate Wiener functional $F = (F^1, \dots, F^n) \in \mathbf{D}_{\infty}(\mathbb{R}^n)$, we denote by $\sigma_F^{ij}(w) = \langle DF^i(w), DF^j(w) \rangle_{\mathcal{H}}$ the (i, j) -component of Malliavin covariance matrix. We denote by $\gamma_F^{ij}(w)$ the (i, j) -component of the inverse matrix σ_F^{-1} . Note that $\sigma_F^{ij} \in \mathbf{D}_{\infty}$ and $D\gamma_F^{ij} = \sum_{k,l} \gamma_F^{ik}(D\sigma_F^{kl})\gamma_F^{lj}$. Hence, derivatives of γ_F^{ij} can be written in terms of γ_F^{ij} 's and the derivatives of σ_F^{ij} 's. Suppose $G \in \mathbf{D}_{\infty}$ and $T \in \mathcal{S}'(\mathbb{R}^n)$. Then, the following integration by parts holds;

$$\mathbb{E}[(\partial_i T \circ F) \cdot G] = \mathbb{E}[(T \circ F) \cdot \Phi_i(\cdot; G)] \quad (4.1)$$

where $\Phi_i(w; G) \in \mathbf{D}_{\infty}$ is given by

$$\begin{aligned} \Phi_i(w; G) = & - \sum_{j=1}^d \left\{ - \sum_{k,l=1}^d G(w) \gamma_F^{ik}(w) \gamma_F^{jl}(w) \langle D\sigma_F^{kl}(w), DF^j(w) \rangle_{\mathcal{H}} \right. \\ & \left. + \gamma_F^{ij}(w) \langle DG(w), DF^j(w) \rangle_{\mathcal{H}} + \gamma_F^{ij}(w) G(w) LF^j(w) \right\}. \end{aligned} \quad (4.2)$$

Note that the expectations in (4.1) are in fact the generalized ones, i.e., the pairing of $\tilde{\mathbf{D}}_{-\infty}$ and $\tilde{\mathbf{D}}_{\infty}$.

Let us recall Watanabe's asymptotic expansion theorem. Let $\{F_{\varepsilon}\}_{0 < \varepsilon \leq 1}$ be a family of \mathbb{R}^n -valued Wiener functionals indexed by a small parameter $\varepsilon \in (0, 1]$. If $\{F_{\varepsilon}\}$ admits an asymptotic expansion in $\mathbf{D}_{\infty}(\mathbb{R}^n)$ and their Malliavin covariance matrices are uniformly non-degenerate, then $T \circ F_{\varepsilon}$ admits an asymptotic expansion in $\tilde{\mathbf{D}}_{-\infty}$ as $\varepsilon \searrow 0$ and each term in the expansion is obtained by formal Taylor expansion. (Theorem 9.4, [11])

In this paper, however, we do not use this method. Instead, we use a modified version of Watanabe's asymptotic expansion theorem, which can be found in pp. 216–217, Takano and Watanabe [25].

Let $\rho > 0$, $\xi \in \mathbf{D}_{\infty}$ and $F \in \mathbf{D}_{\infty}(\mathbb{R}^n)$ and suppose that

$$\inf_{v \in \mathbb{S}^{n-1}} v^* \sigma_F v \geq \rho \quad \text{on} \quad \{w \in \mathcal{W} \mid |\xi(w)| \leq 2\}. \quad (4.3)$$

Let $\chi : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth function whose support is contained in $[-1, 1]$. Then, the following proposition holds (Proposition 6.1, [25]).

Proposition 4.1 *Assume (4.3). For every $T \in \mathcal{S}'(\mathbb{R}^n)$, $\chi(\xi) \cdot T \circ F = \chi(\xi) \cdot T(F) \in \tilde{\mathbf{D}}_{-\infty}$ can be defined in a unique way so that the following properties hold:*

- (i) *If $T_k \rightarrow T \in \mathcal{S}'(\mathbb{R}^n)$ as $k \rightarrow \infty$, then $\chi(\xi) \cdot T_k(F) \rightarrow \chi(\xi) \cdot T(F) \in \tilde{\mathbf{D}}_{-\infty}$.*
- (ii) *If T is given by $g \in \mathcal{S}(\mathbb{R}^n)$, then $\chi(\xi) \cdot T(F) = \chi(\xi)g(F) \in \mathbf{D}_{\infty}$.*

Next, we state the asymptotic expansion theorem, which is Proposition 6.2, [25]. Let $\{F_\varepsilon\}_{0 < \varepsilon \leq 1} \subset \mathbf{D}_\infty(\mathbb{R}^n)$ and $\{\xi_\varepsilon\}_{0 < \varepsilon \leq 1} \subset \mathbf{D}_\infty$ be families of Wiener functionals such that the following asymptotics hold:

$$F_\varepsilon \sim f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^n) \text{ as } \varepsilon \searrow 0, \quad (4.4)$$

$$\xi_\varepsilon \sim a_0 + \varepsilon a_1 + \varepsilon^2 a_2 + \cdots \quad \text{in } \mathbf{D}_\infty \text{ as } \varepsilon \searrow 0. \quad (4.5)$$

Proposition 4.2 *Assume (4.4), (4.5) and $|a_0| \leq 1/8$. Moreover, assume that there exists $\rho > 0$ independent of ε such that (4.3) with $F = F_\varepsilon$ and $\xi = \xi_\varepsilon$ holds for any $\varepsilon \in (0, 1]$. Let $\chi : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth function whose support is contained in $[-1, 1]$ such that $\chi(x) = 1$ if $|x| \leq 1/2$. Then, we have the following asymptotic expansion:*

$$\chi(\xi_\varepsilon) \cdot T(F_\varepsilon) \sim \Phi_0 + \varepsilon \Phi_1 + \varepsilon^2 \Phi_2 + \cdots \quad \text{in } \tilde{\mathbf{D}}_{-\infty} \text{ as } \varepsilon \searrow 0.$$

In the above proposition, $\Phi_k \in \tilde{\mathbf{D}}_{-\infty}$ can be written as the k th coefficient of the formal Taylor expansion of $T(f_0 + [\varepsilon f_1 + \varepsilon^2 f_2 + \cdots])$. In particular, $\Phi_0 = T(f_0)$. (In this paper, however, we do not need the expansion up to high order.)

4.2 Preliminaries from rough path theory

In this subsection we recall the geometric rough path space with Hölder or Besov norm and quasi-sure property of rough path lift. For basic properties of geometric rough path space, we refer to Lyons, Caruana, and Lévy [20], and Friz and Victoir [9]. For the geometric rough path space with Besov norm, we refer to Appendix A.2, [9]. Quasi-sure property of rough path lift is summarized in Inahama [15]. In this paper we basically assume $\alpha \in (1/3, 1/2)$ unless otherwise stated. We always assume that Besov parameters $(\alpha, 4m)$ satisfy (2.3), although some results presented in this subsection still hold under weaker assumptions on the parameters.

We denote by $G\Omega_\alpha^H(\mathbb{R}^d)$ the geometric rough path space over \mathbb{R}^d with α -Hölder norm. For $\beta \in (0, 1]$, let $C_0^{\beta-H}([0, 1], \mathbb{R}^k)$ be the Banach space of all the \mathbb{R}^k -valued, β -Hölder continuous paths that start at 0. If $\alpha + \beta > 1$, then the Young pairing

$$G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{\beta-H}([0, 1], \mathbb{R}^k) \ni (\mathbf{w}, \lambda) \mapsto (\mathbf{w}, \boldsymbol{\lambda}) \in G\Omega_\alpha^H(\mathbb{R}^{d+k})$$

is a well-defined, locally Lipschitz continuous map. (See Section 9.4, [9] for instance.)

Now we consider a system of RDEs driven by the Young pairing $(\mathbf{w}, \boldsymbol{\lambda}) \in G\Omega_\alpha^H(\mathbb{R}^{d+1})$ of $\mathbf{w} \in G\Omega_\alpha^H(\mathbb{R}^d)$ and $\lambda \in C_0^{1-H}([0, 1], \mathbb{R}^1)$. (In most cases, we will assume $\lambda_t = \text{const} \times t$.) For vector fields $V_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($0 \leq i \leq d$), consider

$$dx_t = \sum_{i=1}^d V_i(x_t) dw_t^i + V_0(x_t) d\lambda_t \quad \text{with } x_0 = x \in \mathbb{R}^n. \quad (4.6)$$

The RDEs for the Jacobian process and its inverse are given as follows;

$$dJ_t = \sum_{i=1}^d \nabla V_i(x_t) J_t dw_t^i + \nabla V_0(x_t) J_t d\lambda_t \quad \text{with } J_0 = \text{Id} \in \text{Mat}(n, n), \quad (4.7)$$

$$dK_t = - \sum_{i=1}^d K_t \nabla V_i(x_t) dw_t^i - K_t \nabla V_0(x_t) d\lambda_t \quad \text{with } K_0 = \text{Id} \in \text{Mat}(n, n). \quad (4.8)$$

Here, J, K , and ∇V_i are $\text{Mat}(n, n)$ -valued.

Assume that V_i 's are of C_b^4 for a while. Then, a global solution of (4.6)–(4.8) exists for any \mathbf{x} and λ . Moreover, Lyons' continuity theorem holds. (The linear growth case is complicated and will be discussed later). In that case, the following map is continuous:

$$G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1) \ni (\mathbf{w}, \lambda) \mapsto (\mathbf{x}, \mathbf{J}, \mathbf{K}) \in G\Omega_\alpha^H(\mathbb{R}^n \oplus \text{Mat}(n, n)^{\oplus 2}).$$

(The map $(\mathbf{w}, \lambda) \mapsto \mathbf{x}$ will be denoted by $\Phi : G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1) \rightarrow G\Omega_\alpha^H(\mathbb{R}^n)$.) Recall that in Lyons' formulation of rough path theory, the initial values of the first level paths must be adjusted. When $w \in C_0^{1-H}([0, 1], \mathbb{R}^d)$ and \mathbf{w} is its natural lift, then the path

$$t \mapsto (x + \mathbf{x}_{0,t}^1, \text{Id} + \mathbf{J}_{0,t}^1, \text{Id} + \mathbf{K}_{0,t}^1)$$

is identical to the solution of a system (4.6)–(4.8) of ODEs understood in the Riemann-Stieltjes sense. Recall also that $(\text{Id} + \mathbf{J}_{0,t}^1)^{-1} = \text{Id} + \mathbf{K}_{0,t}^1$ always holds.

We define a continuous function $\Gamma : G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1) \rightarrow \text{Mat}(n, n)$ as follows: Set

$$\Gamma(\mathbf{w}, \lambda) = (\text{Id} + \mathbf{J}_{0,t}^1) C(\mathbf{w}, \lambda) (\text{Id} + \mathbf{J}_{0,t}^1)^*,$$

where

$$C(\mathbf{w}, \lambda) := \int_0^1 (\text{Id} + \mathbf{K}_{0,t}^1) \mathbf{V}(x + \mathbf{x}_{0,t}^1) \mathbf{V}(x + \mathbf{x}_{0,t}^1)^* (\text{Id} + \mathbf{K}_{0,t}^1)^* dt$$

with $\mathbf{V} := [V_1, \dots, V_d] \in \text{Mat}(n, d)$.

If $\lambda_t = t$, then $\Gamma(\varepsilon \mathbf{w}, \varepsilon^2 \lambda) = \sigma_{X_1^\varepsilon}(w)$ for μ -a.a. w , where \mathbf{w} denotes the usual Brownian rough path under μ and X_1^ε denotes the solution of SDE (2.1) at $t = 1$. If $\lambda_t \equiv 0$ and $\mathbf{h} = \mathcal{L}(h)$ is the natural lift of $h \in \mathcal{H}$, then $\Gamma(\mathbf{h}, 0) = \sigma_{\phi_1}(h)$, the deterministic Malliavin covariance matrix given in (3.4).

Remark 4.3 *In this paper we will use Lyons' continuity theorem only with respect to α -Hölder topology ($1/3 < \alpha < 1/2$) and for C_b^4 -coefficient vector fields. We do not try to extend it to the case of unbounded coefficient vector fields or Besov topology.*

For $(\alpha, 4m)$ which satisfies (2.3), $G\Omega_{\alpha, 4m}^B(\mathbb{R}^d)$ denotes the geometric rough path space over \mathbb{R}^d with $(\alpha, 4m)$ -Besov norm. Recall that the distance on this space is given by

$$\begin{aligned} d(\mathbf{w}, \hat{\mathbf{w}}) &= \|\mathbf{w}^1 - \hat{\mathbf{w}}^1\|_{\alpha, 4m-B} + \|\mathbf{w}^2 - \hat{\mathbf{w}}^2\|_{2\alpha, 2m-B} \\ &:= \left(\iint_{0 \leq s < t \leq 1} \frac{|\mathbf{w}_{s,t}^1 - \hat{\mathbf{w}}_{s,t}^1|^{4m}}{|t-s|^{1+4m\alpha}} ds dt \right)^{1/4m} + \left(\iint_{0 \leq s < t \leq 1} \frac{|\mathbf{w}_{s,t}^2 - \hat{\mathbf{w}}_{s,t}^2|^{2m}}{|t-s|^{1+4m\alpha}} ds dt \right)^{1/2m}. \end{aligned}$$

By the Besov-Hölder embedding theorem for rough path spaces, there is a continuous embedding $G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \hookrightarrow G\Omega_{\alpha-(1/4m)}^H(\mathbb{R}^d)$. If $\alpha < \alpha' < 1/2$, there is a continuous embedding $G\Omega_{\alpha'}^H(\mathbb{R}^d) \hookrightarrow G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. We remark that we will not write the first embedding explicitly. (For example, if we write $\Phi(\mathbf{w}, \lambda)$ for $(\mathbf{w}, \lambda) \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1)$, then it is actually the composition of the first embedding map above and Φ with respect to $\{\alpha - 1/(4m)\}$ -Hölder topology.)

Note also that the Young translation by $h \in \mathcal{H}$ works well on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ under (2.3). The map $(\mathbf{w}, h) \mapsto \tau_h(\mathbf{w})$ is continuous from $G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \times \mathcal{H}$ to $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$, where $\tau_h(\mathbf{w})$ is the Young translation of \mathbf{w} by h . (In the proof of the continuity, the only non-trivial components are the "cross integrals" in the second level paths of $\tau_h(\mathbf{w})$, which are computed in Lemma 5.1, [15].)

Now we discuss quasi-sure properties of rough path lift map \mathcal{L} from \mathcal{W} to $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. For $k = 1, 2, \dots$ and $w \in \mathcal{W}$, we denote by $w(k)$ the k th dyadic piecewise linear approximation of w associated with the partition $\{l2^{-k} \mid 0 \leq l \leq 2^k\}$ of $[0, 1]$. We set

$$\mathcal{Z}_{\alpha,4m} := \{w \in \mathcal{W} \mid \{\mathcal{L}(w(k))\}_{k=1}^\infty \text{ is Cauchy in } G\Omega_{\alpha,4m}^B(\mathbb{R}^d)\}.$$

We define $\mathcal{L} : \mathcal{W} \rightarrow G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ by $\mathcal{L}(w) = \lim_{m \rightarrow \infty} \mathcal{L}(w(k))$ if $w \in \mathcal{Z}_{\alpha,4m}$ and we do not define $\mathcal{L}(w)$ if $w \notin \mathcal{Z}_{\alpha,4m}$. (We will always use this version of \mathcal{L} .) Note that \mathcal{H} and $C_0^{\beta-H}([0, 1], \mathbb{R}^d)$ with $\beta \in (1/2, 1]$ are subsets of $\mathcal{Z}_{\alpha,4m}$ and the two definition of rough path lift coincide. Under scalar multiplication and Cameron-Martin translation, $\mathcal{Z}_{\alpha,4m}$ is invariant. Moreover, $c\mathcal{L}(w) = \mathcal{L}(cw)$ and $\tau_h(\mathcal{L}(w)) = \mathcal{L}(w + h)$ for any $w \in \mathcal{Z}_{\alpha,4m}$, $c \in \mathbf{R}$, and $h \in \mathcal{H}$.

It is known that $\mathcal{Z}_{\alpha,4m}^c$ is slim, that is (p, r) -capacity of this set is zero for any $p \in (1, \infty)$ and $r \in \mathbf{N}$. (See Aida [1], Inahama [12, 15]). Therefore, from a viewpoint of quasi-sure analysis, the lift map \mathcal{L} is well-defined. (Quasi-sure property of the lift map is recently extended to the case of a certain class of Gaussian processes by Boediharjo, Geng, and Qian [6].) Moreover, the map $\mathcal{W} \ni w \mapsto \mathcal{L}(w) \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ is ∞ -quasi-continuous (Aida [1]). We will sometimes write $\mathbf{W} := \mathcal{L}(w)$ when it is regarded as a rough path space-valued random variable defined on \mathcal{W} . Due to Lyons' continuity theorem and uniqueness of quasi-continuous modification, $\tilde{X}^\varepsilon(\cdot, x, w) = x + \Phi(\varepsilon\mathcal{L}(w), \varepsilon^2\lambda)^1$ holds quasi-surely, if V_i ($0 \leq i \leq d$) is of C_b^3 . (Here, $\lambda_t = t$.)

Before closing this subsection, we give a brief remark for the coefficient vector fields with linear growth. Below we assume that V_i ($0 \leq i \leq d$) satisfies **(A1)** and has linear growth. In this case it is not easy to prove the existence of a global solution of RDE (4.6) for any $(\mathbf{w}, \lambda) \in G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1)$. (It must be unique if it exists.) Hence, the Lyons-Itô map Φ may not be defined on the whole space $G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1)$.

However, if a global solution \mathbf{x} exists for (\mathbf{w}, λ) , then we can prove with a cut-off technique that a global solution exists for (\mathbf{w}', λ') sufficiently near (\mathbf{w}, λ) , too. Hence, Φ can be defined on an open subset of $G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1)$ and is continuous on it. Let O be the largest open subset with such a property. Then, $\mathcal{L}(\mathcal{H}) \times \{0\} \subset O$.

By Wong-Zakai's approximation theorem (a.s. convergence with respect to the sup-norm will do), we can see that, for each $\varepsilon \in (0, 1]$, $X^\varepsilon(\cdot, x, w) = x + \Phi(\varepsilon\mathcal{L}(w), \varepsilon^2\lambda)^1$,

μ -a.s. on $\{w \in \mathcal{W} \mid (\varepsilon \mathcal{L}(w), \varepsilon^2 \lambda) \in O\}$. Note that the same remark goes even if $G\Omega_\alpha^H(\mathbb{R}^d)$ is replaced by $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. (In this paragraph, $\lambda_t = t$.)

5 Lower estimate

The aim of this section is to prove the lower estimate in our main theorem (Theorem 2.1, (i)). The proof here is more difficult than the one for the elliptic case in the author's previous paper [15]. The keys of the proof are Propositions 3.1 and 4.2.

In what follows, Φ stands for the Lyons-Itô map associated with the vector fields $\{V_1, \dots, V_d; V_0\}$ with respect to $\{\alpha - 1/(4m)\}$ -Hölder topology. We write $\lambda_t^\varepsilon = \varepsilon^2 t$. If $\mathbf{h} := \mathcal{L}(h)$ is the natural lift of $h \in \mathcal{H}$, then $x + \Phi(\mathbf{h}, 0)_{0,t}^1 = \phi_t(h)$, where ϕ is defined by (2.2) and 0 is the constant one-dimensional path 0.

Let X^ε be as in (2.1). It is known that, for any $h \in \mathcal{H}$, $X^\varepsilon(1, x, w + (h/\varepsilon))$ admits an asymptotic expansion in \mathbf{D}_∞ -topology as $\varepsilon \searrow 0$;

$$X^\varepsilon(1, x, w + \frac{h}{\varepsilon}) \sim f_0(h) + \varepsilon f_1(w; h) + \varepsilon^2 f_2(w; h) + \dots \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^n).$$

Here, $f_0(h) = \phi_1(h) = \phi(1, x, h)$ and $f_1(w; h) = D\phi_1(h)\langle w \rangle$, which is continuous, linear in w . (See Section 5, [25]). We do not need precise information of f_i ($i \geq 2$) in this paper. Obviously, $Y^\varepsilon(1, x, w + (h/\varepsilon))$ also admits an asymptotic expansion;

$$Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}) \sim g_0(h) + \varepsilon g_1(w; h) + \varepsilon^2 g_2(w; h) + \dots \quad \text{in } \mathbf{D}_\infty(\mathbb{R}^l).$$

Here, we set $g_i = \Pi_\nu f_i$. In particular, $g_0(h) = \psi_1(h) = \psi(1, x, h)$. Hence, $g_0(h) = a$ if $h \in \mathcal{M}^{x,a}$. Note that $f_1(w; h)$ and $g_1(w; h)$ are mean-zero Gaussian random vectors whose covariance matrices are $\sigma_{\phi_1}(h)$ and $\sigma_{\psi_1}(h)$, respectively.

Let $U \subset G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ be open and \mathcal{L} be the rough path lift map. It suffices to show that

$$\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(U) \geq -\frac{1}{2} \|h\|_{\mathcal{H}}^2 \quad (5.1)$$

for any $h \in \mathcal{M}^{x,a}$ such that $\mathbf{h} \in U$. Here, $\mu_{x,a}^\varepsilon$ is the push-forward by $\varepsilon \mathcal{L} : \mathcal{W} \rightarrow G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ of the finite Borel measure $\theta_{x,a}^\varepsilon$, where $\theta_{x,a}^\varepsilon$ corresponds to the positive Watanabe distribution $\delta_a(Y_1^\varepsilon)$. Moreover, due to Proposition 3.1, it suffices to show (5.1) under the additional conditions on h , namely, (i) $\sigma_{\phi_1}(h)$ is non-degenerate and (ii) $\langle h, \cdot \rangle$ extends to a continuous linear functional on \mathcal{W} .

For $R > 0$, we set

$$\hat{B}_R = \{\mathbf{w} \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \mid \|\mathbf{w}^1\|_{\alpha,4m-B} + \|\mathbf{w}^2\|_{2\alpha,2m-B}^{1/2} < R\}$$

and set $\hat{B}_R(\mathbf{h}) = \tau_h(\hat{B}_R)$, where τ_h is the Young translation by h on $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. Since τ_h is a homeomorphism for any $h \in \mathcal{H}$, $\{\hat{B}_R(\mathbf{h}) \mid R > 0\}$ forms a fundamental system of

open neighborhood around \mathbf{h} . Since U is open, $\hat{B}_R(\mathbf{h}) \subset U$ for sufficiently small $R > 0$. We will estimate the weight of $\hat{B}_R(\mathbf{h})$ from below instead of that of U .

Let $\rho > 0$ be such that $v^* \Gamma(\mathbf{h}, 0)v = v^* \sigma_{\phi_1(h)} v \geq 2\rho$ for any $v \in \mathbb{S}^{n-1}$. Since Γ is continuous from $G\Omega_{\alpha, 4m}^B(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}^1)$, there exist $R_0 > 0$ and $\varepsilon_0 \in (0, 1]$ such that

$$v^* \Gamma(\tau_h(\mathbf{w}), \lambda^\varepsilon) v \geq \rho \quad (v \in \mathbb{S}^{n-1}, \mathbf{w} \in \hat{B}_{R_0}, \varepsilon \in [0, \varepsilon_0]).$$

Even when V_i is of linear growth, the left hand side is well-defined for small enough $R > 0$. Note that the Malliavin covariance matrix of

$$\varepsilon^{-1} X^\varepsilon(1, x, w + (h/\varepsilon)) = \varepsilon^{-1} \{x + \Phi(\tau_h(\varepsilon \mathbf{W}), \lambda^\varepsilon)^1\}$$

equals $\Gamma(\tau_h(\varepsilon \mathbf{W}), \lambda^\varepsilon)$ a.s. Recall that the smallest eigenvalue of the Malliavin covariance matrix of $\varepsilon^{-1} Y^\varepsilon(1, x, w + (h/\varepsilon))$ is larger than or equal to that of $\varepsilon^{-1} X^\varepsilon(1, x, w + (h/\varepsilon))$. Hence, the smallest eigenvalue of the Malliavin covariance matrix of $\varepsilon^{-1} Y^\varepsilon(1, x, w + (h/\varepsilon)) \geq \rho$, provided that $\varepsilon \mathbf{W} \in \hat{B}_{R_0}$ and $\varepsilon \in [0, \varepsilon_0]$.

By Cameron-Martin formula, it holds that, for any $F \in \mathbf{D}_\infty$,

$$\begin{aligned} \mathbb{E}[F \delta_a(Y_1^\varepsilon)] &= \mathbb{E}[\exp(-\frac{\langle h, w \rangle}{\varepsilon} - \frac{\|h\|_{\mathcal{H}}^2}{2\varepsilon^2}) F(w + \frac{h}{\varepsilon}) \delta_a(Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}))] \\ &= e^{-\|h\|_{\mathcal{H}}^2/2\varepsilon^2} \varepsilon^{-l} \mathbb{E}[e^{-\langle h, w \rangle/\varepsilon} F(w + \frac{h}{\varepsilon}) \delta_0(\varepsilon^{-1}[Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}) - a])] \end{aligned}$$

Here, we used the fact that $\delta_0(\varepsilon \cdot) = \varepsilon^{-l} \delta_0(\cdot)$.

Let $\chi : \mathbb{R} \rightarrow \mathbb{R}$ be as in Proposition 4.2. Moreover, we assume that χ is non-increasing on $[0, \infty)$ so that χ takes values in $[0, 1]$. For sufficiently small $R > 0$, we have

$$\begin{aligned} \mu_{x,a}^\varepsilon(\hat{B}_R(\mathbf{h})) &= \int I_{\hat{B}_R(\mathbf{h})}(\mathbf{w}) \mu_{x,a}^\varepsilon(d\mathbf{w}) = \int I_{\hat{B}_R}(\tau_{-h}(\mathbf{w})) \mu_{x,a}^\varepsilon(d\mathbf{w}) \\ &= \int I_{\hat{B}_R}(\tau_{-h}(\varepsilon \mathbf{W})) \theta_{x,a}^\varepsilon(dw) \\ &\geq \int \chi\left(\frac{\|\tau_{-h}(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B}^{4m} + \|\tau_{-h}(\varepsilon \mathbf{W})^2\|_{2\alpha, 2m-B}^{2m}}{R^{4m}}\right) \theta_{x,a}^\varepsilon(dw) \\ &= \mathbb{E}\left[\chi\left(\frac{\|\tau_{-h}(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B}^{4m} + \|\tau_{-h}(\varepsilon \mathbf{W})^2\|_{2\alpha, 2m-B}^{2m}}{R^{4m}}\right) \delta_a(Y_1^\varepsilon)\right] \\ &= e^{-\|h\|_{\mathcal{H}}^2/2\varepsilon^2} \varepsilon^{-l} \mathbb{E}\left[e^{-\langle h, w \rangle/\varepsilon} \chi\left(\frac{\|(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B}^{4m} + \|(\varepsilon \mathbf{W})^2\|_{2\alpha, 2m-B}^{2m}}{R^{4m}}\right) \right. \\ &\quad \left. \times \delta_0\left(\varepsilon^{-1}[Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}) - a]\right)\right]. \end{aligned}$$

Note that $w = \mathbf{W}^1$ and if $\|(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B} \leq R$, then $\|w\|_\infty \leq c_1 R/\varepsilon$, where $c_1 > 0$ is the operator norm of the embedding of the usual path space with $(\alpha, 4m)$ -Besov norm into the one with the sup-norm. Therefore, $e^{-\langle h, w \rangle/\varepsilon} \geq \exp(-c_2 R/\varepsilon^2)$, where $c_2 := c_1 \|\langle h, \cdot \rangle\|_{\mathcal{W}^*}$ is a positive constant independent of R, ε . Noting that the positive Watanabe distribution

$\delta_0(\varepsilon^{-1}[Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}) - a])$ is in fact a finite measure on \mathcal{W} by Sugita's theorem [24], we see that

$$\mu_{x,a}^\varepsilon(\hat{B}_R(\mathbf{h})) \geq e^{-\|h\|_{\mathcal{H}}^2/2\varepsilon^2} \varepsilon^{-l} e^{-c_2 R/\varepsilon^2} \times \mathbb{E} \left[\chi \left(\frac{\|(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B}^{4m} + \|(\varepsilon \mathbf{W})^2\|_{2\alpha, 2m-B}^{2m}}{R^{4m}} \right) \delta_0 \left(\varepsilon^{-1}[Y^\varepsilon(1, x, w + \frac{h}{\varepsilon}) - a] \right) \right].$$

Assume that the logarithm of the generalized expectation on the right hand side above converges to some (finite) real quantity (for each $R > 0$ small enough) as $\varepsilon \searrow 0$, which we will prove later. Then, we have

$$\liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(\hat{B}_{R_0}(\mathbf{h})) \geq \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(\hat{B}_R(\mathbf{h})) \geq -\frac{1}{2} \|h\|_{\mathcal{H}}^2 - c_2 R$$

if $R \in (0, R_0)$. Letting $R \searrow 0$, we obtain the desired estimate (5.1).

Now we use Proposition 4.2 with $\xi_\varepsilon = \{\|(\varepsilon \mathbf{W})^1\|_{\alpha, 4m-B}^{4m} + \|(\varepsilon \mathbf{W})^2\|_{2\alpha, 2m-B}^{2m}\}/R^{4m}$, $T = \delta_0$, and $F_\varepsilon = \varepsilon^{-1}[Y^\varepsilon(1, x, w + (h/\varepsilon)) - a]$. Notice that if $0 < R \leq 2^{-1/m} R_0$, then the condition (4.3) is satisfied with $F = F_\varepsilon$, $\xi = \xi_\varepsilon$, and $\rho > 0$ defined as above. Proposition 4.2 implies that

$$\chi(\xi_\varepsilon) \cdot \delta_0(F_\varepsilon) = \delta_0(g_1(w; h)) + O(\varepsilon) \quad \text{in } \tilde{\mathbf{D}}_{-\infty} \text{ as } \varepsilon \searrow 0$$

Since $g_1(w; h)$ is a non-degenerate Gaussian random variable taking values in \mathbb{R}^l , its law has a strictly positive density. Hence, we have

$$\lim_{\varepsilon \searrow 0} \mathbb{E}[\chi(\xi_\varepsilon) \cdot \delta_0(F_\varepsilon)] = \mathbb{E}[\delta_0(g_1(w; h))] \in (0, \infty).$$

Therefore, $\lim_{\varepsilon \searrow 0} \log \mathbb{E}[\chi(\xi_\varepsilon) \cdot \delta_0(F_\varepsilon)] \in (-\infty, \infty)$, which completes the proof of the lower estimate of our main theorem.

Remark 5.1 *In the last part of the proof above, we used a modified version of Watanabe's asymptotic expansion (Proposition 4.2). However, as in the proof of the elliptic case in [15], it may also be nice to use the standard version (Theorem 9.4, [11]) after proving the uniform non-degeneracy of $\varepsilon^{-1}[Y^\varepsilon(1, x, w + (h/\varepsilon)) - a]$ when the deterministic Malliavin covariance matrix at h is non-degenerate. (Loosely speaking, the authors of [4] argue in this way, for instance.)*

In the hypoelliptic case, however, the proof of uniform non-degeneracy becomes more difficult. We need to combine (i) Kusuoka-Stroock's bound (6.2) below for Y^ε instead of X^ε and (ii) the Schilder-type LDP for Brownian rough path.

6 Upper estimate

The aim of this section is to prove the upper estimate in our main theorem (Theorem 2.1, (i)). The proof here is similar to the one for the elliptic case in [15] and is a modification of it.

In this section, we will often use the following fact; For $f, g : (0, 1] \rightarrow [0, \infty)$, it holds that $\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log(A_\varepsilon + B_\varepsilon) \leq [\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log A_\varepsilon] \vee [\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log B_\varepsilon]$. We assume without loss of generality that $\mathcal{V} = \mathbb{R}^l \times \{\mathbf{0}_{n-l}\}$ so that $Y_t^\varepsilon = \Pi_{\mathcal{V}}(X_t^\varepsilon) = (X_t^{\varepsilon,1}, \dots, X_t^{\varepsilon,l})$. (This assumption is just for notational simplicity.)

[Step 1] We divide the proof into three steps. The first step is to show that

$$\lim_{R \searrow 0} \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(B_R(\mathbf{w})) \leq -I(\mathbf{w}), \quad \mathbf{w} \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d), \quad (6.1)$$

where

$$B_R(\mathbf{w}) = \{\mathbf{v} \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \mid \|\mathbf{v}^i - \mathbf{w}^i\|_{i\alpha,4m/i-B} < R^i \quad (i = 1, 2)\}.$$

First, we consider the case $\mathbf{w} = \mathbf{h}$, where $h \in \mathcal{H} \setminus \mathcal{M}^{x,a}$. We write $\tilde{a} := \Pi_{\mathcal{V}}(x + \Phi(\mathbf{h}, 0)_{0,1}^1) (\neq a)$. By the arguments in the previous section, even if V_i 's admits linear growth, there exist $\varepsilon_0 > 0$ and $R > 0$ such that $(\mathbf{v}, \varepsilon) \mapsto \Phi(\mathbf{v}, \lambda^\varepsilon)$ is well-defined and continuous on $B_{2^{1/4m}R}(\mathbf{h}) \times [0, \varepsilon_0]$. Moreover, we may assume that $|\Pi_{\mathcal{V}}(x + \Phi(\mathbf{v}, \lambda^\varepsilon)_{0,1}^1) - \tilde{a}| \leq |a - \tilde{a}|/3$ holds for all $0 \leq \varepsilon \leq \varepsilon_0$ and $\mathbf{v} \in B_{2^{1/4m}R}(\mathbf{h})$. Note that for a fixed $\varepsilon \in [0, \varepsilon_0]$, $X^\varepsilon(\cdot, x, w) = x + \Phi(\varepsilon \mathbf{W}, \lambda^\varepsilon)^1$ holds μ -a.s. on $\{w \in \mathcal{W} \mid \varepsilon \mathbf{W} \in B_{2^{1/4m}R}(\mathbf{h})\}$.

Let us verify that $\mu_{x,a}^\varepsilon(B_R(\mathbf{h})) = 0$ as follows. Then, (6.1) immediately follows. Let $\chi : \mathbb{R} \rightarrow [0, 1]$ be a smooth even function such that $\chi = 1$ on $[0, 1]$ and $\chi = 0$ on $[2, \infty)$ and non-increasing on $[1, 2]$. Let $\{f_k\}$ be a sequence of continuous functions on \mathbb{R}^l such that $f_k \rightarrow \delta_a$ in $\mathcal{S}'(\mathbb{R}^l)$ as $k \rightarrow \infty$. We may assume that the support of f_k is contained in $\{\xi \in \mathbb{R}^l \mid |\xi - a| < |a - \tilde{a}|/3\}$ for any k .

Then we have,

$$\begin{aligned} \mu_{x,a}^\varepsilon(B_R(\mathbf{h})) &= \theta_{x,a}^\varepsilon(\{w \in \mathcal{W} \mid \varepsilon \mathbf{W} \in B_R(\mathbf{h})\}) \\ &\leq \int_{\mathcal{W}} \prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{h}^i\|_{i\alpha,4m/i-B}^{4m/i} / R^{4m}) \theta_{x,a}^\varepsilon(dw) \\ &= \mathbb{E} \left[\prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{h}^i\|_{i\alpha,4m/i-B}^{4m/i} / R^{4m}) \delta_a(Y_1^\varepsilon) \right] \\ &= \lim_{k \rightarrow \infty} \mathbb{E} \left[\prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{h}^i\|_{i\alpha,4m/i-B}^{4m/i} / R^{4m}) f_k(Y_1^\varepsilon) \right] \\ &= \lim_{k \rightarrow \infty} \mathbb{E} \left[\prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{h}^i\|_{i\alpha,4m/i-B}^{4m/i} / R^{4m}) f_k(\Pi_{\mathcal{V}}(x + \Phi(\varepsilon \mathbf{W}, \lambda^\varepsilon)_{0,1}^1)) \right] = 0. \end{aligned}$$

Next, let us consider the other case, namely $\mathbf{w} \in G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \setminus \mathcal{L}(\mathcal{H} \setminus \mathcal{M}^{x,a})$. Note that $\|D^r X_1^\varepsilon\|_{L^p}$ is bounded in ε for any $p \in (1, \infty)$ and $r = 0, 1, 2, \dots$, where D stands for the \mathcal{H} -derivative. Recall that the Malliavin covariance matrix of X_1^ε satisfies the following estimate;

$$\|(\det \sigma_{X_1^\varepsilon})^{-1}\|_{L^p} \leq K_1(p) \varepsilon^{-K_2} \quad \text{for all } p \in (1, \infty) \text{ and } \varepsilon \in (0, 1]. \quad (6.2)$$

Here, the constant $K_1(p) > 0$ may depend on p but not on ε and the constant $K_2 > 0$ does not depend on p, ε (the starting point x is fixed here). This can be found in Section V-10, [11] or originally in Kusuoka-Stroock [18].

Note that $\|\varepsilon \mathbf{W}^1 - \mathbf{w}^1\|_{\alpha, 4m-B}^{4m}$ and $\|\varepsilon^2 \mathbf{W}^2 - \mathbf{w}^2\|_{2\alpha, 2m-B}^{2m}$ belong to the $4m$ -th order inhomogeneous Wiener chaos. Since their L^2 -norms are bounded in ε , so are their (p, r) -Sobolev norms for any (p, r) .

Set $G(u_1, \dots, u_l) = \prod_{j=1}^l (u_j - a_j)^+$, which is a continuous function from \mathbb{R}^l to \mathbb{R} with polynomial growth and satisfies $\partial_1^2 \cdots \partial_l^2 G(u) = \delta_a(u)$ in the sense of Schwartz distributions on \mathbb{R}^l . It is straight forward to check that $(\partial_1^2 \cdots \partial_l^2 G) \circ \Pi_V = (\partial_1^2 \cdots \partial_l^2)(G \circ \Pi_V)$ as a Schwartz distribution on \mathbb{R}^n .

Then, we have

$$\begin{aligned} \mu_{x,a}^\varepsilon(B_R(\mathbf{w})) &= \theta_{x,a}^\varepsilon(\{w \in \mathcal{W} \mid \varepsilon \mathbf{W} \in B_R(\mathbf{w})\}) \\ &\leq \mathbb{E} \left[\prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{w}^i\|_{i\alpha, 4m/i-B}^{4m/i} / R^{4m}) \cdot (\partial_1^2 \cdots \partial_l^2 G)(Y_1^\varepsilon) \right] \\ &\leq \mathbb{E} \left[\prod_{i=1}^2 \chi(\|\varepsilon^i \mathbf{W}^i - \mathbf{w}^i\|_{i\alpha, 4m/i-B}^{4m/i} / R^{4m}) \cdot (\partial_1^2 \cdots \partial_l^2)(G \circ \Pi_V)(X_1^\varepsilon) \right]. \end{aligned} \quad (6.3)$$

Now we use the integration by parts formula for Watanabe distributions as in p. 377, [11]. Then, the right hand side of (6.3) is equal to a finite sum of the following form;

$$\sum_{j,j'} \mathbb{E} \left[F_{j,j'}^\varepsilon \cdot \chi^{(j)} \left(\frac{\|\varepsilon \mathbf{W}^1 - \mathbf{w}^1\|_{\alpha, 4m-B}^{4m}}{R^{4m}} \right) \chi^{(j')} \left(\frac{\|\varepsilon^2 \mathbf{W}^2 - \mathbf{w}^2\|_{2\alpha, 2m-B}^{2m}}{R^{4m}} \right) (G \circ \Pi_V)(X_1^\varepsilon) \right]. \quad (6.4)$$

Here, $F_{j,j'}^\varepsilon(w) = F_{j,j'}^\varepsilon(\varepsilon, w)$ is a polynomial in components of (i) X_1^ε and its derivatives, (ii) $\|\varepsilon^i \mathbf{W}^i - \mathbf{w}^i\|_{i\alpha, 4m/i-B}^{4m/i}$ and its derivatives, (iii) $\sigma_{X_1^\varepsilon}$, which is a Malliavin covariance matrix of X_1^ε , and (iv) $\gamma_{X_1^\varepsilon} = (\sigma_{X_1^\varepsilon})^{-1}$. Note that derivatives of $\gamma_{X_1^\varepsilon}$ do not appear. It is important that the right hand side of (6.4) is not a generalized expectation anymore. By (6.2) there exists a constant $K > 0$ such that $F_{j,j'}^\varepsilon$ is $O(\varepsilon^{-K})$ in any L^p . (Below K may change from line to line. The exact value of K is of no importance.)

Take any $p, q \in (1, \infty)$ such that $1/p + 1/q = 1$. By Hölder's inequality, the right hand side of (6.4) is dominated by

$$C \varepsilon^{-K} \mu \left(\|\varepsilon^i \mathbf{W}^i - \mathbf{w}^i\|_{i\alpha, 4m/i-B}^{1/i} \leq 2^{\frac{1}{4m}} R \ (i = 1, 2) \right)^{\frac{1}{q}} = C \varepsilon^{-K} \mu(\varepsilon \mathbf{W} \in B_{2^{1/4m} R}(\mathbf{w}))^{\frac{1}{q}}.$$

Here, $C = C_{p,q} > 0$ is a constant independent of ε . By the large deviation principle of Schilder-type for the scaled Brownian rough path $\varepsilon \mathbf{W}$ on $G\Omega_{\alpha, 4m}^B(\mathbb{R}^d)$, we have

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(B_R(\mathbf{w})) \leq -\frac{1}{q} \inf \{ \|h\|_{\mathcal{H}}^2 / 2 \mid h \in \mathcal{H}, \mathcal{L}(h) \in B_{2^{1/4m} R}(\mathbf{w}) \}.$$

Letting $q \searrow 1$, we have

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(B_R(\mathbf{w})) &\leq -\inf\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{H}, \mathcal{L}(h) \in B_{2^{1/4m}R}(\mathbf{w})\} \\ &= -\inf\{I_{Sch}(\mathbf{v}) \mid \mathbf{v} \in B_{2^{1/4m}R}(\mathbf{w})\}. \end{aligned}$$

Since the good rate function $I_{Sch} : G\Omega_{\alpha,4m}^B(\mathbb{R}^d) \rightarrow [0, \infty]$ is lower semicontinuous, the limit of the right hand side as $R \searrow 0$ is dominated by $-I_{Sch}(\mathbf{w})$. This proves (6.1). (Here, $I_{Sch}(\mathcal{L}(h)) := \|h\|_{\mathcal{H}}^2/2$ and $I_{Sch}(\mathbf{w}) := \infty$ if \mathbf{w} is not the natural lift of any $h \in \mathcal{H}$.)

[Step 2] The second step is to prove the upper bound in our main theorem (Theorem 2.1, (i)) when A is a compact set in $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. Let $N \in \mathbb{N}$ be sufficiently large. For any $\mathbf{w} \in A$, take $R = R_{N,\mathbf{w}} > 0$ small enough so that

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(B_R(\mathbf{w})) \leq \begin{cases} -N & (\text{if } I(\mathbf{w}) = \infty), \\ -I(\mathbf{w}) + N^{-1} & (\text{if } I(\mathbf{w}) < \infty). \end{cases}$$

The union of such open balls over $\mathbf{w} \in A$ covers the compact set A . Hence, there are finitely many $\mathbf{w}_1, \dots, \mathbf{w}_k \in A$ such that $A \subset \cup_{i=1}^k B_{R_i}(\mathbf{w}_i)$, where $R_i = R(N, \mathbf{w}_i)$. By using the remark in the beginning of this section, we see that

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) &\leq (-N) \vee \max\{-I(\mathbf{w}_i) + N^{-1} \mid 1 \leq i \leq k, I(\mathbf{w}_i) < \infty\} \\ &\leq (-N) \vee \left(-\inf_{h \in \mathcal{L}^{-1}(A) \cap \mathcal{M}^{x,a}} \|h\|_{\mathcal{H}}^2/2 + N^{-1}\right). \end{aligned}$$

Letting $N \rightarrow \infty$, we obtain

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) \leq -\inf\{\|h\|_{\mathcal{H}}^2/2 \mid h \in \mathcal{M}^{x,a}, \mathcal{L}(h) \in A\}.$$

Thus, we have obtained the upper estimate for the compact case.

[Step 3] In this final step we will prove the upper bound in our main theorem (Theorem 2.1, (i)) when A is a closed set in $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$.

Take α' slightly greater than α so that the condition (2.3) still holds for (α', m) . Then, the continuous embedding $G\Omega_{\alpha',4m}^B(\mathbb{R}^d) \hookrightarrow G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$ is in fact compact, which means that any bounded set in $G\Omega_{\alpha',4m}^B(\mathbb{R}^d)$ is precompact in $G\Omega_{\alpha,4m}^B(\mathbb{R}^d)$. (See [15].)

For $\rho > 0$, denote by $B'_\rho(\mathbf{0})$ the ball with respect to $(\alpha', 4m)$ -Besov norm of radius ρ and centered at the trivial rough path $\mathbf{0}$. Then, $B'_\rho(\mathbf{0})$ is precompact with respect to $(\alpha, 4m)$ -Besov topology.

Then, $\overline{A \cap B'_\rho(\mathbf{0})}$ is compact with respect to $(\alpha, 4m)$ -Besov topology and is included in $A = \bar{A}$, where the closure is taken with respect to $(\alpha, 4m)$ -Besov topology. Hence, we can use the argument in the previous step:

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(\overline{A \cap B'_\rho(\mathbf{0})}) &\leq -\inf\left\{\frac{\|h\|_{\mathcal{H}}^2}{2} \mid h \in \mathcal{M}^{x,a}, \mathcal{L}(h) \in \overline{A \cap B'_\rho(\mathbf{0})}\right\} \\ &\leq -\inf\left\{\frac{\|h\|_{\mathcal{H}}^2}{2} \mid h \in \mathcal{M}^{x,a}, \mathcal{L}(h) \in A\right\}. \end{aligned}$$

On the other hand, the weight of $B'_\rho(\mathbf{0})^c$ is dominated as follows.

$$\begin{aligned}
\mu_{x,a}^\varepsilon(B'_\rho(\mathbf{0})^c) &= \theta_{x,a}^\varepsilon(\{w \in \mathcal{W} \mid \varepsilon \mathbf{W} \notin B'_\rho(\mathbf{0})\}) \\
&\leq \|\delta_a(Y_1^\varepsilon)\|_{2,-r} \text{Cap}_{2,r}(\{w \in \mathcal{W} \mid \varepsilon \mathbf{W} \notin B'_\rho(\mathbf{0})\}) \\
&\leq \|\delta_a(Y_1^\varepsilon)\|_{2,-r} \left[\text{Cap}_{2,r}(\{w \in \mathcal{W} \mid \|\mathbf{W}^1\|_{\alpha',4m-B} \geq \frac{\rho}{\varepsilon}\}) \right. \\
&\quad \left. + \text{Cap}_{2,r}(\{w \in \mathcal{W} \mid \|\mathbf{W}^2\|_{2\alpha',2m-B}^{1/2} \geq \frac{\rho}{\varepsilon}\}) \right]. \quad (6.5)
\end{aligned}$$

Here r is a sufficiently large integer and $\text{Cap}_{2,r}$ is the capacity associated with $\mathbf{D}_{2,r}$. Recall that $\theta_{x,a}^\varepsilon$ is associated with $\delta_a(Y_1^\varepsilon)$ via Sugita's theorem.

By the large deviation estimate for capacities in [15], the second factor on the right hand side of (6.5) is known to be dominated by $\exp(-c(\rho/\varepsilon)^2)$ when ρ/ε is sufficiently large. Here, $c = c(\alpha', m, 2, r)$ is a positive constant.

Suppose that

$$\|\delta_a(Y_1^\varepsilon)\|_{2,-r} = O(\varepsilon^{-\nu}) \quad \text{as } \varepsilon \searrow 0 \quad (6.6)$$

holds for some $r \in \mathbb{N}$ and some $\nu > 0$. Then, from (6.5) and (6.6), we see that

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A \cap B'_\rho(\mathbf{0})^c) = \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(B'_\rho(\mathbf{0})^c) \leq -c\rho^2.$$

Hence,

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) \leq \left(-\inf \left\{ \frac{\|h\|_{\mathcal{H}}^2}{2} \mid h \in \mathcal{M}^{x,a}, \mathcal{L}(h) \in A \right\} \right) \vee (-c\rho^2).$$

Letting $\rho \rightarrow \infty$, we have

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu_{x,a}^\varepsilon(A) \leq -\inf \left\{ \frac{\|h\|_{\mathcal{H}}^2}{2} \mid h \in \mathcal{M}^{x,a}, \mathcal{L}(h) \in A \right\},$$

which is the desired upper estimate.

Now, it remains to prove (6.6). We use the integration by parts formula as in (6.3)–(6.4) in Step 1. However, it is actually easier this time. (The same symbols are used below.) Let $Q \in \mathbf{D}_\infty$ be arbitrary. We have

$$\mathbb{E}[Q \cdot \delta_a(Y_1^\varepsilon)] = \mathbb{E}[Q \cdot (\partial_1^2 \cdots \partial_t^2)(G \circ \Pi_\nu)(X_1^\varepsilon)] = \mathbb{E}[F_Q^\varepsilon \cdot (G \circ \Pi_\nu)(X_1^\varepsilon)].$$

Here, $F_Q^\varepsilon(w)$ is a polynomial in components of (i) X_1^ε and its derivatives, (ii) Q and its derivatives, (iii) $\sigma_{X_1^\varepsilon}$, which is a Malliavin covariance matrix of X_1^ε , and (iv) $\gamma_{X_1^\varepsilon} = (\sigma_{X_1^\varepsilon})^{-1}$. Note that derivatives of $\gamma_{X_1^\varepsilon}$ do not appear. The right hand side of (6.4) is not a generalized expectation anymore.

Note that F_Q is linear in Q and there exists $r \in \mathbb{N}$ such that the order of derivatives of Q which are involved in the expression of F_Q^ε is bounded from above by r . Combining these with (6.2), we see that

$$|\mathbb{E}[Q \cdot \delta_a(Y_1^\varepsilon)]| \leq C\varepsilon^{-\nu} \|Q\|_{2,r} \quad (Q \in \mathbf{D}_\infty)$$

for some $\nu > 0$ and $C > 0$, which are independent of ε and Q . Since \mathbf{D}_∞ is dense in $\mathbf{D}_{2,r}$, we obtain (6.6). This completes the proof of the upper estimate.

7 Proof of Corollary 2.4

In this section we prove Corollary 2.4. Since we are familiar with probability measures, we prove the second assertion, from which the first assertion immediately follows. When the vector fields A_i ($0 \leq i \leq d$) are bounded, we can use Lyons continuity theorem for C_b^3 -vector fields and the proof is quite simple due to the contraction principle. However, when the vector fields have linear growth, we will rely on a cut-off argument and our proof looks a little bit complicated. (In this section we will write $\lambda_t^\varepsilon = \varepsilon^2 t$.)

7.1 Bounded case

Consider the product measure $\hat{\mu}_{x,a}^\varepsilon \otimes \delta_{\lambda^\varepsilon}$ on $G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R})$. Since the second component λ^ε is deterministic and continuous in ε , it follows from (the Hölder version of) Theorem 2.1, (ii) that $\{\hat{\mu}_{x,a}^\varepsilon \otimes \delta_{\lambda^\varepsilon}\}_{0 < \varepsilon \leq 1}$ also satisfies an LDP with a good rate function J . Here, the effective domain of J is $\mathcal{L}(\mathcal{M}^{x,a}) \times \{0\}$ and $J(\mathcal{L}(h), 0) = \|h\|_{\mathcal{H}}^2/2$ for $h \in \mathcal{M}^{x,a}$.

Let $\Phi' : G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R}) \rightarrow G\Omega_\alpha^H(\mathbb{R}^N)$ be the Lyons-Itô map associated with A_i ($0 \leq i \leq d$). Then, $\hat{Z}^\varepsilon(\cdot, z, w) = z + \Phi'(\varepsilon \mathcal{L}(w), \lambda^\varepsilon)^1$ quasi-surely. (See [15].) In particular, they coincide $\hat{\theta}_{x,a}^\varepsilon$ -almost surely. Then, $\tilde{Z}^\varepsilon(\cdot, z)_*[\hat{\theta}_{x,a}^\varepsilon] = (z + (\Phi')^1)_*[\hat{\mu}_{x,a}^\varepsilon \otimes \delta_{\lambda^\varepsilon}]$. Note also that $\zeta(h) = z + \Phi'(\mathcal{L}(h), 0)^1$. Now, by the contraction principle (Theorem 4.2.1, [7]), we can easily show Corollary 2.4, (ii) in this case.

7.2 Linearly growing case

In this case it is not so easy to see whether the Lyons-Itô map Φ' is everywhere-defined continuous map or not. However, as we mentioned before, it is well-defined and continuous around (the lift of) Cameron-Martin space. Hence, we use a cut-off argument and a modified version of the contraction principle (Lemma 7.1 below). This method has already been used for the usual Freidlin-Wentzell type large deviations when the coefficient vector fields admit linear growth (see [13]).

The following lemma is a slight modification of the contraction principle for LDPs and is formulated in a general setting. It states that, if the map is continuous around the effective domain of the good rate function, then the contraction principle still holds. The map need not be continuous everywhere.

Lemma 7.1 *Let S and \hat{S} be polish spaces and let $f : S \rightarrow \hat{S}$ be a measurable map. We assume that $\{\mu_\varepsilon\}_{\varepsilon > 0}$ is a family of probability measures on S which satisfies an LDP with a good rate function J as $\varepsilon \searrow 0$. Let $\mathcal{D} = \{a \in S \mid J(a) < \infty\}$ be the effective domain of J . Assume further that there is an open subset U of S such that $\mathcal{D} \subset U$ and $f|_U$ is continuous. Then, $\{\mu_\varepsilon \circ f^{-1}\}_{\varepsilon > 0}$ satisfies an LDP with a good rate function \hat{J} as $\varepsilon \searrow 0$, where $\hat{J}(b) = \inf\{J(a) \mid a \in f^{-1}(\{b\})\}$.*

Proof. For a proof, see Lemma 2.3 [13], for instance. (One can also prove this lemma by hand, since the proof is not so different from the one for the standard contraction principle.) ■

Now we discuss a refinement of the Wong-Zakai approximation. For $k = 1, 2, \dots$ and $w \in \mathcal{W}$, $w(k)$ stands for the k th dyadic polygonal approximation as before. We consider the following ODE in the Riemann-Stieltjes sense.

$$dz(k)_t^\varepsilon = \varepsilon \sum_{i=1}^d A_i(z(k)_t^\varepsilon) dw(k)_t^i + \varepsilon^2 A_0(z(k)_t^\varepsilon) dt \quad \text{with} \quad z(k)_0^\varepsilon = z \quad (7.1)$$

It is well-known that if A_i ($0 \leq i \leq d$) satisfies **(A1)**, then for each fixed ε it holds that $\lim_{k \rightarrow \infty} \sup_{0 \leq t \leq 1} |Z_t^\varepsilon - z(k)_t^\varepsilon| = 0$ for μ -a.a. w . However, this convergence actually takes place quasi-surely.

In the next lemma O denotes the largest open subset of $G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R})$ on which Φ' is well-defined. As we already explained, Φ' is continuous from O to $G\Omega_\alpha^H(\mathbb{R}^N)$.

Lemma 7.2 *Assume that the vector field A_i satisfies **(A1)** for any $0 \leq i \leq d$ and fix any $\varepsilon \in (0, 1]$. Then, quasi-surely,*

$$\lim_{k \rightarrow \infty} \sup_{0 \leq t \leq 1} |\tilde{Z}_t^\varepsilon - z(k)_t^\varepsilon| = 0.$$

Here, $\tilde{Z}^\varepsilon = \tilde{Z}^\varepsilon(\cdot, z, \cdot)$ stands for the ∞ -quasi-sure modification of Z^ε defined in (2.6). Moreover, we have

$$\tilde{Z}^\varepsilon(\cdot, z, w) = z + \Phi'(\varepsilon \mathbf{W}, \lambda^\varepsilon)^1 \quad \text{quasi-surely on } \{w \in \mathcal{W} \mid (\varepsilon \mathbf{W}, \lambda^\varepsilon) \in O\}.$$

Proof. For any $0 < a < b$, choose a smooth, non-increasing function $\chi_{a,b} : \mathbb{R} \rightarrow [0, 1]$ such that $\chi_{a,b} = 1$ on $(-\infty, a]$, $\chi_{a,b} > 0$ on $(-\infty, b)$, and $\chi_{a,b} = 0$ on $[b, \infty)$. For $\nu = 1, 2, \dots$, we set $A_i^\nu(z) = \chi_{\nu+1, \nu+2}(|z|) A_i(z)$ for $z \in \mathbb{R}^N$ and $0 \leq i \leq d$. Clearly, A_i^ν is of C_b^∞ and agrees with A_i on $\{z \mid |z| \leq \nu + 1\}$. Consider the scaled SDE (2.4) and its approximating ODE (7.1) with their coefficient vector fields being replaced by A_i^ν . The solutions are denoted by $Z^{\varepsilon, \nu}$ and $z^{\varepsilon, \nu}(k)$, respectively. The Lyons-Itô map associated with the new coefficients is denoted by Φ'_ν , which is defined everywhere and continuous from $G\Omega_\alpha^H(\mathbb{R}^d) \times C_0^{1-H}([0, 1], \mathbb{R})$ to $G\Omega_\alpha^H(\mathbb{R}^N)$. Then, $\tilde{Z}^{\varepsilon, \nu} = z + \Phi'_\nu(\varepsilon \mathbf{W}, \lambda^\varepsilon)^1$, quasi-surely.

Take any ν such that $\nu \geq |z|$ and we will denote the sup-norm by $\|\cdot\|_\infty$. By a standard argument for stopping times, $\{w \mid \|Z^\varepsilon\|_\infty < \nu\} = \{w \mid \|Z^{\varepsilon, \nu}\|_\infty < \nu\}$, μ -a.s. and $Z^\varepsilon = Z^{\varepsilon, \nu}$, μ -a.s. on this subset. Hence, for any $0 < a < b < \nu + 1$, $\chi_{a,b}(\|Z^\varepsilon\|_\infty) Z^\varepsilon = \chi_{a,b}(\|Z^{\varepsilon, \nu}\|_\infty) Z^{\varepsilon, \nu}$, μ -a.s. on \mathcal{W} . By the uniqueness of quasi-continuous modification, $\chi_{a,b}(\|\tilde{Z}^\varepsilon\|_\infty) \tilde{Z}^\varepsilon = \chi_{a,b}(\|\tilde{Z}^{\varepsilon, \nu}\|_\infty) \tilde{Z}^{\varepsilon, \nu}$, quasi-surely on \mathcal{W} . Assume that $z \neq 0$, since the case $z = 0$ can be shown with trivial modification. Since \tilde{Z}^ε and $\tilde{Z}^{\varepsilon, \nu}$ can never be a

zero path, this implies that $\{w \mid \|Z^\varepsilon\|_\infty < b\} = \{w \mid \|Z^{\varepsilon,\nu}\|_\infty < b\}$, quasi-surely, for any $b \in (0, \nu + 1)$. Taking $b = a$ and using the above equality once again, we have $Z^\varepsilon = Z^{\varepsilon,\nu}$ quasi-surely on the above subset for any $b \in (0, \nu + 1)$. (Below, we will use this fact with $b = \nu$.)

If $w \in \{w \mid \|Z^\varepsilon\|_\infty < \nu\} = \{w \mid \|Z^{\varepsilon,\nu}\|_\infty < \nu\}$ and admits a rough path lift with respect to α -Hölder rough path topology, then it is easy to see that

$$Z^\varepsilon(\cdot, z, w) = Z^{\varepsilon,\nu}(\cdot, z, w) = z + \Phi'_\nu(\varepsilon\mathcal{L}(w), \lambda^\varepsilon)^1 = \lim_{m \rightarrow \infty} \left(z + \Phi'_\nu(\varepsilon\mathcal{L}(w(k)), \lambda^\varepsilon)^1 \right),$$

where we have used Lyons' continuity theorem for Φ'_ν . Since $z + \Phi'_\nu(\varepsilon\mathcal{L}(w(k)), \lambda^\varepsilon)^1 = z^{\varepsilon,\nu}(k)(\cdot, z, w)$ stays inside the ball of radius $\nu + 1$ for sufficiently large k , it holds that $z^{\varepsilon,\nu}(k)(\cdot, z, w) = z^\varepsilon(k)(\cdot, z, w)$. Thus, we have shown the refinement of the Wong-Zakai approximation on the set $\{w \mid \|Z^\varepsilon\|_\infty < \nu\}$ and, by taking the union with respect to ν , on the whole Wiener space, too.

The proof of the second assertion of the lemma is quite similar. We just need to note that Φ' is continuous on the open set O and that $z + \Phi'_\nu(\varepsilon\mathcal{L}(w), \lambda^\varepsilon)^1 = z + \Phi'(\varepsilon\mathcal{L}(w), \lambda^\varepsilon)^1$ as long as it stays inside the ball of radius $\nu + 1$ for sufficiently large k . \blacksquare

Proof of the linear growth case of Corollary 2.4. For simplicity of notation we prove the case $z = 0$ only. We extend $(\Phi')^1 : O \rightarrow C^{\alpha-H}([0, 1], \mathbb{R}^N)$ by setting $\Phi'(\mathbf{w}, \lambda)^1 = 0$ if $(\mathbf{w}, \lambda) \notin O$. Note that O contains $\mathcal{L}(\mathcal{H}) \times \{0\}$, which in turn contains the effective domain of the rate function \hat{I}_1 in Theorem 2.1. Then, by Lemma 7.1, the push-forward measure of $\hat{\mu}_{x,a}^\varepsilon$ by the map $(\Phi')^1$ satisfies an LDP with a good rate function \hat{I}_2 .

Fix $\varepsilon \in (0, 1]$. On the probability space $(\mathcal{W}, \hat{\theta}_{x,a}^\varepsilon)$, we have two $C^{\alpha-H}([0, 1], \mathbb{R}^N)$ -valued random maps. One is \tilde{Z}^ε . The other is $w \mapsto \Phi'(\varepsilon\mathbf{W}, \lambda^\varepsilon)^1$. The push-forward measure of $\hat{\theta}_{x,a}^\varepsilon$ by the latter map is $(\Phi')^1_*[\hat{\mu}_{x,a}^\varepsilon]$, which we have just discussed.

Let us consider the set on which these two maps disagree. By Lemma 7.2 and the fact that $\hat{\theta}_{x,a}^\varepsilon$ does not charge a slim set,

$$\begin{aligned} \hat{\theta}_{x,a}^\varepsilon(\{w \mid \tilde{Z}^\varepsilon(\cdot, z, w) \neq \Phi'(\varepsilon\mathbf{W}, \lambda^\varepsilon)^1\}) &\leq \hat{\theta}_{x,a}^\varepsilon(\{w \mid (\varepsilon\mathbf{W}, \lambda^\varepsilon) \notin O\}) \\ &= \hat{\mu}_{x,a}^\varepsilon \otimes \delta_{\lambda^\varepsilon}(O^c). \end{aligned}$$

Our main theorem (Theorem 2.1) implies that $\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \hat{\mu}_{x,a}^\varepsilon \otimes \delta_{\lambda^\varepsilon}(O^c) = -\infty$. Therefore, these two random maps are exponentially equivalent in the sense of Definition 4.2.10, [7]. We see from Theorem 4.2.13, [7] that $(\tilde{Z}^\varepsilon)_*[\hat{\theta}_{x,a}^\varepsilon]$ also satisfies an LDP with the same good rate function \hat{I}_2 . This completes the proof of Corollary 2.4, (ii). \blacksquare

References

- [1] Aida, S.; Vanishing of one-dimensional L^2 -cohomologies of loop groups. J. Funct. Anal. 261 (2011), no. 8, 2164–2213.

- [2] Aida, S.; Kusuoka, S.; Stroock, D.; On the support of Wiener functionals. Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990), 3–34, Pitman Res. Notes Math. Ser., 284, Longman Sci. Tech., Harlow, 1993.
- [3] Bailleul, I., Large deviation principle for bridges of degenerate diffusion processes. Preprint (2013). Arxiv Math: 1303.2854.
- [4] Bailleul, I., Mesnager, L., Norris, J. R.; Small time fluctuations for bridges of Riemannian and sub-Riemannian diffusions. Preprint (2014).
- [5] Ben Arous, G.; Léandre, R.; Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. Probab. Theory Related Fields 90 (1991), no. 3, 377–402.
- [6] Boedihardjo, H.; Geng, X.; Qian, Z.; Quasi-sure convergence theorem in p-variation distance for Gaussian sample paths. Preprint. arXiv:1309.0835.
- [7] Dembo, A., Zeitouni, O.; Large deviations techniques and applications. Second edition. Springer-Verlag, New York, 1998.
- [8] Friz, P.; Victoir, N.; Large deviation principle for enhanced Gaussian processes. Ann. Inst. H. Poincaré Probab. Statist. 43 (2007), no. 6, 775–785.
- [9] Friz, P.; Victoir, N.; Multidimensional stochastic processes as rough paths. Cambridge University Press, Cambridge, 2010.
- [10] Hsu, P.; Brownian bridges on Riemannian manifolds. Probab. Theory Related Fields 84 (1990), no. 1, 103–118.
- [11] Ikeda, N., Watanabe, S.; Stochastic differential equations and diffusion processes. Second edition. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989.
- [12] Inahama, Y.; Quasi-sure existence of Brownian rough paths and a construction of Brownian pants. Infin. Dimens. Anal. Quantum Probab. Relat. Top. 9 (2006), no. 4, 513–528.
- [13] Inahama, Y.; A note on rough differential equations with unbounded coefficients. Probabilistic approach to geometry, 155–170, Adv. Stud. Pure Math., 57, Math. Soc. Japan, Tokyo, 2010.
- [14] Inahama, Y.; Laplace approximation for rough differential equation driven by fractional Brownian motion, Ann. Probab. 41 (2013), No. 1, 170–205.
- [15] Inahama, Y.; Large deviation principle of Freidlin-Wentzell type for pinned diffusion processes. To appear in Trans. Amer. Math. Soc. (2014). arXiv:1203.5177.

- [16] Inahama, Y.; Short time kernel asymptotics for rough differential equation driven by fractional Brownian motion, preprint, (2014). arXiv:1403.3181
- [17] Kunita, H.; Supports of diffusion processes and controllability problems. Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), pp. 163–185, Wiley, New York-Chichester-Brisbane, 1978.
- [18] Kusuoka, S., Stroock, D. W.; Applications of the Malliavin calculus. II. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), no. 1, 1–76.
- [19] Ledoux, M.; Qian, Z.; Zhang, T.; Large deviations and support theorem for diffusion processes via rough paths. Stochastic Process. Appl. 102 (2002), no. 2, 265–283.
- [20] Lyons, T.; Caruana, M.; Lévy, T.; Differential equations driven by rough paths. Lecture Notes in Math., 1908. Springer, Berlin, 2007.
- [21] Malliavin, P.; Stochastic analysis. Springer-Verlag, Berlin, 1997.
- [22] Nualart, D.; The Malliavin calculus and related topics. Second edition. Springer-Verlag, Berlin, 2006.
- [23] Shigekawa, I.; Stochastic analysis. Translations of Mathematical Monographs, 224. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 2004.
- [24] Sugita, H.; Positive generalized Wiener functions and potential theory over abstract Wiener spaces. Osaka J. Math. 25 (1988), no. 3, 665–696.
- [25] Takanobu, S.; Watanabe, S.; Asymptotic expansion formulas of the Schilder type for a class of conditional Wiener functional integrations. Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990), 194–241, Pitman Res. Notes Math. Ser., 284, Longman Sci. Tech., Harlow, 1993.

Yuzuru INAHAMA

Graduate School of Mathematics, Nagoya University,
Furocho, Chikusa-ku, Nagoya 464-8602, JAPAN.

Email: inahama@math.nagoya-u.ac.jp